

Inflationary perturbation theory is geometrical optics in phase space

David Seery,^{1,*} David J. Mulryne,^{2,†} Jonathan Frazer,^{1,‡} and Raquel H. Ribeiro^{3,§}

¹*Astronomy Centre, University of Sussex, Brighton BN1 9QH, United Kingdom*

²*Astronomy Unit, School of Mathematical Sciences, Queen Mary, University of London, Mile End Road, London E1 4NS, United Kingdom*

³*Department of Applied Mathematics and Theoretical Physics, Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, United Kingdom*

A pressing problem in comparing inflationary models with observation is the accurate calculation of correlation functions. One approach is to evolve them using ordinary differential equations (“transport equations”), analogous to the Schwinger–Dyson hierarchy of *in-out* quantum field theory. We extend this approach to the complete set of momentum space correlation functions. A formal solution can be obtained using raytracing techniques adapted from geometrical optics. We reformulate inflationary perturbation theory in this language, and show that raytracing reproduces the familiar “ δN ” Taylor expansion. Our method produces ordinary differential equations which allow the Taylor coefficients to be computed efficiently. We use raytracing methods to express the gauge transformation between field fluctuations and the curvature perturbation, ζ , in geometrical terms. Using these results we give a compact expression for the nonlinear gauge-transform part of f_{NL} in terms of the principal curvatures of uniform energy-density hypersurfaces in field space.

I. INTRODUCTION

Our current theories of the early universe are stochastic. They do not predict a definite *state* today: rather, their predictions are statistical. To compare these predictions with observation it must usually be supposed that we are in some sense typical. This brings two challenges. First, what is typical under some circumstances may be atypical under others. Therefore we must be precise about the type of observer of which we are a typical representative. This leads to the “measure problem,” about which we have nothing new to say. In this paper we are concerned with the second challenge: after fixing a class of observers, to estimate the observables typically measured by its members.

Inflation is the most common early-universe paradigm for which we would like to compute observables. In this context we usually take ourselves to be ordinary observers of the fluctuations produced on approach to a fixed vacuum. The challenge is to calculate the typical stochastic properties of these fluctuations.

The most important fluctuation generated by inflation is the primordial density perturbation, ζ . Correlations in the temperature and polarization anisotropies of the microwave background are inherited from ζ and provide a clean probe of its statistical character. Therefore, both present-day constraints [1] and the imminent arrival of high quality microwave-background data [2] make accurate estimates of its statistical properties a pressing issue. Meanwhile, large surveys of the cosmological density field will provide information about its properties on complementary, smaller scales [3]. To compare this abundance of

data to models we require an efficient tool with which to estimate the n -point functions $\langle \zeta^n \rangle$.

Taking ζ to be synthesized from the fluctuation of one or more light scalar fields during an inflationary era, several computational schemes exist which enable the n -point functions to be studied. Many of these schemes employ some variant of the *separate universe picture* [4]. Taking H to be the Hubble parameter, this asserts that—when smoothed on some physical scale, L , much larger than the horizon scale, so that $L/H^{-1} \gg 1$ —the average evolution of each L -sized patch can be computed using the *background* equations of motion and initial conditions taken from smoothed quantities local to the patch. Working from a Taylor expansion in the initial conditions for each patch, Lyth & Rodríguez showed how this assumption could be turned into a practical algorithm for calculating correlation functions [5]. This “ δN method” has become the most popular way to explore the predictions of specific models, both analytically and numerically, and has developed a large literature of its own. The principal difficulty arises when calculating the coefficients of the Taylor expansion, sometimes called the “ δN coefficients.” We shall discuss this difficulty in §IVB.

Alternative approaches exist. Rigopoulos, Shellard & van Tent [6] evolved each correlation function using a Langevin equation. Yokoyama, Suyama & Tanaka [7] decomposed each δN coefficient into components which could be computed using ordinary differential equations. Later, a systematic method to obtain ‘transport’ equations for the entire hierarchy of correlation functions (rather than simply the δN coefficients) was introduced [8]. A more longstanding approach uses the methods of traditional cosmological perturbation theory (“CPT”) to produce “transfer matrices” [9]. This has recently been revived by a number of authors [10–14]. Numerical approaches have been employed by Lehnert & Renaux-Petel [15], Ringeval [16], and Huston & Malik [17].

* D.Seery@sussex.ac.uk

† D.Mulryne@qmul.ac.uk

‡ J.Frazer@sussex.ac.uk

§ R.Ribeiro@damtp.cam.ac.uk

The relationship of these different methods to each other has not always been clear. Nor is it always obvious how to relate the approximations employed by each technique. In this paper we study the connections between many of these approaches using the formalism of Elliston et al. [18]. This is a statistical interpretation of the separate universe picture. In what follows we briefly summarize the construction. (See also Ref. [8].)

The separate universe approximation as statistical mechanics.—Fix a large spacetime box of comoving side μ containing the region of interest. The scale μ should be much larger than the separate universe scale, requiring $\mu \gg L$, but not superexponentially larger [19]. After smoothing on the scale L , the fields within the large box pick out an ensemble or cloud of $N \sim (\mu/L)^3$ points in the classical phase space. The condition that μ/L is not *superexponentially* large means that the typical diameter of the cloud will be roughly of order the quantum scatter $\langle \delta \phi^2 \rangle^{1/2} \sim H$. Because N is still large, $N \gg 1$, it is convenient to describe the ensemble by an occupation probability ρ on phase space.¹ The correlation functions of ζ on the scale L are then determined by the classical statistical mechanics of this ensemble, which is encoded in the Boltzmann equation.

In familiar applications of statistical mechanics, the evolution of the ensemble may be complicated. Small-scale interactions scatter members of the cloud between orbits on phase space, represented by the collisional term in the Boltzmann equation. However, the separate universe assumption requires causality to suppress those interactions which would be required for scattering between orbits. Therefore the evolution is trivial. Each point in phase space is assigned an occupation probability by the initial conditions, which is conserved along its orbit. All that is required is a mapping of initial conditions to the final state, which is obtained by carrying the initial conditions along the phase space flow generated by the underlying theory. It follows that the Boltzmann equation can be integrated using the method of characteristics.

A similar conclusion applies to the correlation functions of interest, $\langle \zeta^n \rangle$. These probe information about the distribution function over the cloud, giving a weighted average over many characteristics. Alternatively, if the cloud has only a small phase-space diameter, we can exchange information about the entire set of characteristics for the details of a *single* fiducial characteristic and a description of how nearby characteristics separate from it. In differential geometry this description is provided by the apparatus of Ja-

cobi fields; see Fig. 1. We shall see that the differing implementations of the separate universe approximation can be understood as alternative methods to compute these Jacobi fields.

In applications we are frequently interested in correlation functions associated with mixed scales, rather than a single scale L . To do so we construct multiple ensembles associated with different smoothing scales. The separate universe approximation couples the evolution of all these ensembles in a specific way, which we describe in §III.

Outline.—In this paper we develop and refine the statistical-mechanical interpretation of the separate universe picture summarized above. Because the final distribution of occupation probabilities is an *image* generated by dragging along the phase space flow, it can be calculated in precisely the same way that geometrical optics enables us to calculate the image generated by a source of light rays. In §II we show that, at least within the slow-roll approximation, this parallel is exact; the scalar field equation can be interpreted as the eikonal equation for a light ray in a medium with varying refractive index—or equivalently as Huygens’ equation for a wavefront.

In §§II C–II D we introduce the idea of Jacobi fields and explore their connection with the “adiabatic limit,” in which all isocurvature modes decay and the curvature perturbation becomes conserved. Such limits are important because an inflationary model is predictive *on its own* only if the flow enters such a region [18, 20–22]. Jacobi fields are familiar from the description of congruences of light rays in general relativity [23, 24]. In this case, as shown in Fig. 1, they describe fluctuations between the L -sized patches which make up the ensemble. Their evolution enables the adiabatic and isocurvature modes to be tracked. In particular, decay of isocurvature modes means decay of the corresponding Jacobi fields, which occurs when the bundle of trajectories undergoes focusing.

In §III we use these ideas to develop evolution (“transport”) equations for each correlation function, and in §IV we show that the Jacobi fields can be used to formally integrate the system of transport equations. This gives a practical method to identify regions where the flow becomes adiabatic. The analysis can begin from either the separate universe principle or traditional cosmological perturbation theory. As a by-product, our formal solution demonstrates that the transport equations are equivalent to the Taylor expansion algorithm introduced by Lyth & Rodríguez.

In §IV B we use this solution to derive a closed set of differential equations for the Taylor coefficients, and in §IV C we explain how the transport equations can be manipulated to obtain evolution equations for the coefficients of each momentum “shape”. These shapes will be an important diagnostic tool when comparing inflationary models to observation [25]. Together with the transport hierarchy of §III, the equations of §§IV B–IV C represent the principal results of this paper. Either set can be used to obtain the correlation functions of a given theory, and we discuss their comparative advantages.

In §IV D we give more a more general discussion of the

¹ To be certain that we are estimating only the observables measured by a typical observer living within a single terminal vacuum, we should demand that ρ has support only on points whose orbits eventually converge in some neighbourhood of that vacuum. This requires that all horizon volumes reheat almost surely in the same minimum. If some horizon volumes reheat in different minima then the resulting correlation functions are not measurable by a local observer who sees only a single vacuum.

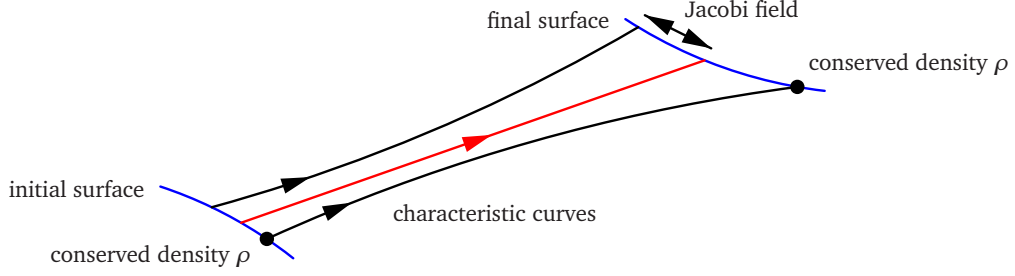


FIG. 1: Jacobi fields. Characteristic curves are labelled by arrows, and the red characteristic is the fiducial curve. A conserved probability density is dragged along the flow. At any point, the Jacobi fields span the space infinitesimal displacements to neighbouring characteristics.

relationship between the transport equations and other formulations of perturbation theory.

In §V we specialize to the slow-roll approximation and use ray-tracing methods to derive the gauge transformation between field fluctuations and the curvature perturbation, ζ . As a result, we obtain the gauge transformation in terms of geometrical quantities—in particular, the extrinsic curvature of constant density hypersurfaces. We separate the gauge contribution to f_{NL} into a number of effects, corresponding to these geometrical quantities. For some models, we show that the largest of these can be attributed to a strong *relative* enhancement of the power in isocurvature fluctuations. We briefly discuss what conclusions can be drawn regarding the asymptotic magnitude of $|f_{\text{NL}}|$.

Finally, we provide a brief summary of our results in §VI.

Notation and conventions.—We use units in which $c = \hbar = 1$, and work in terms of the reduced Planck mass, $M_{\text{P}}^{-2} = 8\pi G$. We use a number of index conventions which are introduced in the text. See especially the paragraph *Index convention* on p. 5, and the discussion of primed indices below Eq. (31) on p. 10.

II. GEOMETRICAL OPTICS IN PHASE SPACE

Throughout this paper, our discussion will apply to an inflationary phase which can be described by a collection of canonical scalar fields ϕ_α coupled to Einstein gravity. We initially use Greek labels α, β, \dots , to label the different species of fields. The action for this system is

$$S = \frac{1}{2} \int d^4x \sqrt{-g} (M_{\text{P}}^2 R - \partial_\alpha \phi_\alpha \partial^\alpha \phi_\alpha - 2V), \quad (1)$$

where $V = V(\phi^\alpha)$ is an interaction potential depending only on the scalar fields, and indices a, b, \dots , run over space time dimensions. We take the background geometry to be flat Friedmann–Robertson–Walker with scale factor $a(t)$.

A. Slow-roll approximation: rays on field space

In this subsection we impose the slow-roll approximation. This requires $\epsilon = -\dot{H}/H^2 \ll 1$ where $H = \dot{a}/a$ is the

Hubble parameter. We introduce an ϵ -parameter for each species of light field,

$$\epsilon_\alpha \equiv \frac{1}{2M_{\text{P}}^2} \frac{\dot{\phi}_\alpha^2}{H^2}, \quad (2)$$

in terms of which one can write $\epsilon = \sum_\alpha \epsilon_\alpha$. The slow-roll approximation therefore entails $\epsilon_\alpha \ll 1$.

Huygens' equation.—Combining (2) and the field equation for ϕ_α , and making use of the slow-roll approximation, we find

$$\frac{d\phi_\alpha}{dN} = \pm M_{\text{P}} \sqrt{2\epsilon_\alpha} = -M_{\text{P}}^2 \partial_\alpha \ln V, \quad (3)$$

where $dN \equiv d \ln a$ measures the number of e-foldings of expansion experienced along the trajectory, and ∂_α denotes a partial derivative with respect to ϕ_α . Eq. (3) constrains the canonical momenta $\sim \dot{\phi}_\alpha$ to lie on a submanifold of the classical phase space coordinatized by the fields ϕ_α . This simplification is a consequence of the slow-roll approximation. In a theory with M scalar fields, it implies that we may work with the simpler M -dimensional field space instead of the full $2M$ -dimensional phase space. This is convenient, although when we later abandon the slow-roll approximation we will have to return to phase space.

In what follows we often rewrite (3) in the form

$$\frac{d\phi_\alpha}{dN} = u_\alpha, \quad \text{where } u_\alpha \equiv -M_{\text{P}}^2 \partial_\alpha f \quad \text{and } f \equiv \ln \frac{V}{V_*}, \quad (4)$$

and interpret the solution $\phi_\alpha(N)$ as an integral curve of the vector field u_α , parametrized by N . The scale V_* is arbitrary. Since f is a gradient, these integral curves correspond to pure potential flow.²

The unit vector parallel to u_α is

$$\hat{n}_\alpha \equiv \frac{u_\alpha}{M_{\text{P}} v}, \quad (5)$$

where we have defined

$$v \equiv \sqrt{2\epsilon}. \quad (6)$$

² Since $df/dN = u_\alpha \partial_\alpha f = -\|u_\alpha\|^2/M_{\text{P}}^2$, it follows that f is monotone decreasing along each integral curve. Therefore one may loosely think of f as a Lyapunov function (or Morse function) for the flow.

It follows that the arc length along an integral curve, labelled s and measured using a flat Euclidean metric on field space, satisfies $ds = M_p v dN$. Reparametrizing each curve in terms of s , the flow equation (4) can be rewritten

$$v \frac{d\phi_\alpha}{ds} = -M_p \partial_\alpha \ln V = \partial_\alpha S, \quad (7)$$

where $S \equiv -M_p f$ is Hamilton's characteristic function. Eq. (7) is *Huygens' equation*. Under the assumptions of geometrical optics, it describes the propagation of a light ray in a medium of spatially varying refractive index v .

Snell's law.—We conclude that the inflationary trajectories in field space are precisely the light rays of geometrical optics, for which Huygens' equation can be thought of as a generalization of the Ibn Sahl or Snell–Descartes law. The wavefronts correspond to level sets of the characteristic function S and are therefore equipotentials, or surfaces of constant energy density in field space. Each light ray is locally orthogonal to these surfaces, so the vector \hat{n}_α is locally the unit vector normal to a surface of constant energy density.

When slow-roll is a good approximation v is small, $v \ll 1$, and increases to $v \sim \mathcal{O}(1)$ near the end of inflation.

B. Rays on phase space

In some circumstances the slow-roll approximation is not available. This may be the case during inflation if slow-roll is temporarily violated—perhaps during a turn in field space, to be studied in §VB—or on approach to the end of inflation, where $\epsilon \sim 1$.

In such cases we must return to the full second-order field equation, which cannot be written in the form of Eq. (3). To obtain an analogue of geometrical optics one must pass to a Hamiltonian formalism. We define

$$p_\alpha = \frac{d\phi_\alpha}{dN}. \quad (8)$$

This plays the role of Huygens' equation for ϕ_α . In terms of p_α , the scalar field equation becomes

$$\frac{dp_\alpha}{dN} = [\epsilon(p) - 3]p_\alpha - \frac{V_\alpha(\phi)}{H(\phi, p)^2}. \quad (9)$$

We must also rewrite ϵ and H in terms of p_α , obtaining

$$\epsilon(p) \equiv -\frac{\dot{H}}{H^2} = \frac{p_\alpha p_\alpha}{2M_p^2} \quad (10a)$$

$$H(\phi, p)^2 M_p^2 = \frac{V(\phi)}{3 - \epsilon(p)}. \quad (10b)$$

Note that ϵ is purely a function of p_α , whereas H is a function of both ϕ_α and p_α .

Eqs. (8)–(9) show that, beyond slow-roll, the *precise* analogy with Huygens' equation is lost. Although these

equations define a congruence of rays in phase space, it is not possible to find a characteristic function S so that these rays are everywhere orthogonal to equipotentials of S . Such a function would have to satisfy $\partial_{\phi_\alpha} S = p_\alpha$, and therefore $S = p_\alpha \phi_\alpha + g(p)$ for arbitrary g . Unfortunately, there is no choice for g which reproduces the right-hand side of Eq. (9).

The majority of our analysis requires only the first-order evolution equations (8)–(9), and at this level the formalism we develop will apply to evolution in phase space without imposing slow-roll. For that purpose it is convenient to combine ϕ_α and p_α into a single phase-space coordinate. We continue to write this ϕ_α , with the understanding that α now ranges over the $2M$ dimensions of phase space. The velocity vector is likewise u_α .

C. Jacobi fields and beam cross-sections

To proceed, we must carry the initial distribution of occupation probabilities along the flow, forming the “image” distribution of interest. In optical language, our task is to understand how images generated from a source of light rays are distorted by passage through a medium.

It was explained above that the typical spacing between arbitrarily selected members of the ensemble should be roughly of order the quantum scatter, $\sigma \sim \langle \delta \phi^2 \rangle^{1/2}$. Because $\sigma/M_p \sim 10^{-5} \ll 1$, this is small in comparison with the natural scale M_p . Therefore the orbits traversed by the cloud trace out a narrowly-collimated spray or “bundle” of light rays in phase space. In canonical models of inflation, setting initial conditions near horizon-crossing will make the initial profile close to Gaussian [26]. Therefore the evolution of the ensemble is similar to the evolution of tightly-focused Gaussian laser beam propagating in an optical cavity.

Connecting vectors.—Cross-sections within the laser beam may be focused, sheared or rotated by refraction. These possibilities are familiar from the study of weak gravitational lensing.

To obtain a quantitative description we slice the laser beam open, generating a cross-section. The precise slicing is arbitrary. For applications to inflation we will often slice along surfaces of fixed energy density, or after a fixed number of e-folds. Distortions of the cross section can be studied if we know how an arbitrary basis is transported from slice to slice. In general relativity this would be Fermi–Walker transport [23].

Jacobi used this method to study geodesic deviation on Riemannian manifolds. For this reason an infinitesimal vector propagated along the beam is called a *Jacobi field*. Taking $\delta \phi_\alpha$ to be such a field and the flow vector u_α to be sufficiently smooth, it will be transported by the equation

$$\frac{d\delta \phi_\alpha}{dN} = \delta \phi_\beta \partial_\beta u_\alpha = u_{\alpha\beta} \delta \phi_\beta. \quad (11)$$

The quantity $u_{\alpha\beta} \equiv \partial_\beta u_\alpha$ is the *expansion tensor*. It can be expanded in terms of a dilation $\theta = \text{tr} u_{\alpha\beta}$, a traceless

symmetric shear $\sigma_{\alpha\beta}$ and an antisymmetric twist $\omega_{\alpha\beta}$,

$$u_{\alpha\beta} \equiv \frac{\theta}{d} \delta_{\alpha\beta} + \sigma_{\alpha\beta} + \omega_{\alpha\beta}, \quad (12)$$

where $d = M$ for flows on field space, or $d = 2M$ if we do not impose the slow-roll approximation and work on the full phase space. In either case $\delta_{\alpha\beta}$ is the Kronecker δ .

Optical scalars.—Dilation describes rigid, isotropic rescaling of $\delta\phi_\alpha$ by $1 + \theta$. It represents a global tendency of the light rays to focus or defocus. The shear $\sigma_{\alpha\beta}$ is a symmetric square matrix and can therefore be diagonalized, yielding d eigenvalues ξ_i and corresponding eigenvectors $s_{\alpha,i}$ representing the principal shear directions (here i is a label taking values $1, \dots, d$; see §VB). The shear describes a rescaling of the component of the connecting vector in the direction $s_{\alpha,i}$ by a factor $1 + \xi_i$. Tracelessness of $\sigma_{\alpha\beta}$ implies $\sum_i \xi_i = 0$, so expansion in one direction must be accompanied by contraction in another. Therefore shear preserves cross-sectional area. Finally, the twist $\omega_{\alpha\beta}$ describes a rigid volume-preserving rotation of $\delta\phi_\alpha$, representing a tendency of neighbouring trajectories to rotate around each other.

It is useful to define σ^2 to satisfy

$$\sigma^2 \equiv \frac{1}{2} \sigma_{\alpha\beta} \sigma_{\alpha\beta}. \quad (13)$$

Imposing the slow-roll approximation and working on field space, the flow is orthogonal to equipotentials of Hamilton's characteristic function. Therefore it is a pure potential flow, for which $\omega_{\alpha\beta} = 0$. On the full phase space this property is lost and the twist can be non-zero. In such cases it is helpful to define $2\omega^2 = \omega_{\alpha\beta} \omega_{\alpha\beta}$. Together, θ , σ^2 and ω^2 comprise the optical scalars introduced by Sachs and Penrose [27].

van Vleck matrix.—Eq. (11) has a well-known formal solution in terms of an ordered exponential [28]. This method was used Rigopoulos, Shellard & van Tent [29], and later by Yokoyama et al. [7]. It yields an explicit (but formal) expression for transport of any Jacobi field along the beam,

$$\delta\phi_\alpha(N) = \Gamma_{\alpha\beta}(N, N_0) \delta\phi_\beta(N_0), \quad (14)$$

where $\delta\phi_\beta(N_0)$ is the Jacobi field on some initial slice $N = N_0$. Eq. (14) describes the evolution of this Jacobi field at any later time N . The matrix $\Gamma_{\alpha\beta}(N, N_0)$ satisfies

$$\Gamma_{\alpha\beta}(N, N_0) \equiv \mathcal{P} \exp \int_{N_0}^N u_{\alpha\beta}(N') dN', \quad (15)$$

where the path-ordering operator \mathcal{P} rewrites its argument with early times on the right-hand side, and later times on the left. We will occasionally refer to $\Gamma_{\alpha\beta}$ as the propagator matrix. It is closely related to a Wilson line.

Index convention.—Eq. (15) can be simplified with the aid of an index convention. Up to this point we have been labelling field-space indices using Greek symbols α, β , etc. To

avoid writing the time of evaluation explicitly, we adopt the convention that Greek indices denote evaluation at the late time of interest, N . Latin indices i, j , etc., denote evaluation at the early time N_0 . Therefore Γ can be written as a mixed index object, $\Gamma_{\alpha i}$.

Eq. (14) immediately implies

$$\Gamma_{\alpha i} = \frac{\partial \phi_\alpha}{\partial \phi_i}, \quad (16)$$

and endows this derivative with a geometric interpretation. It plays an important role in the Lyth–Rodríguez implementation of the separate universe approximation [5], where it appears due to a Taylor expansion in the initial conditions local to each L -sized patch. In this formulation, one often projects on to equipotential surfaces in field space. We define $h_{\alpha\beta} = \partial \phi_\alpha^c / \partial \phi_\beta$ to obtain

$$\frac{\partial \phi_\alpha^c}{\partial \phi_i} = h_{\alpha\beta} \Gamma_{\beta i}. \quad (17)$$

The notation ‘c’ indicates that $d\phi_\alpha^c$ can be thought of as the variation of a field ϕ_α^c defined on a fixed comoving *spacetime* hypersurface [30, 31]. It follows from geometrical arguments that $h_{\alpha\beta} = \delta_{\alpha\beta} - \hat{n}_\alpha \hat{n}_\beta$, where \hat{n}_α is the unit normal to phase-space slices of constant potential energy, defined in (5). The tensor $h_{\alpha\beta}$ is the induced metric (or “first fundamental form”) on these surfaces. Eq. (17) shows that choice of gauge is associated with projection onto an appropriate hypersurface in phase space. Moreover, Eqs. (16)–(17) show that partial derivatives with respect to ϕ_i are associated with propagation of Jacobi fields along the bundle.

Caustics.—The matrix $\Gamma_{\alpha i}$ appears whenever it is necessary to track the distortion of a line element along a flow, and has applications in fluid dynamics, general relativity and elsewhere [23, 32]. DeWitt–Morette observed that, considered as a matrix of Jacobi fields, Eq. (16) was related to the inverse of the van Vleck matrix, introduced in the construction of semiclassical (“WKB”) approximations to the path integral [33].³ We define

$$\Gamma_{i\alpha}^{-1} = \delta_{i\beta} \mathcal{P} \exp \left(- \int_{N_0}^N u_{\beta\alpha}(N') dN' \right). \quad (18)$$

³ In DeWitt–Morette [33] the proof is ascribed to B.S. DeWitt. DeWitt–Morette noted that the relation between Jacobi fields and variation of a general solution of the field equations with respect to its constants of integration had been known to Jacobi (ultimately leading to his development of what is now Hamilton–Jacobi theory), and suggested that this technique could be used to simplify the long calculations which arise when solving Jacobi's equation. Applied to inflationary correlation functions, the history has been reversed: the variational formulae came first, in the form of the Lyth–Rodríguez algorithm. This often leads to simple analytic results, as DeWitt–Morette foresaw. But, as we explain in §IV B, this method is unsuited to numerical implementation, because of the small numerical tolerances required to reliably determine variation with respect to the initial conditions. It is preferable to solve an *ordinary* differential equation, such as Jacobi's equation (11) or (25).

The van Vleck matrix is $\Delta_{i\alpha} \equiv (N - N_0)^d \Gamma_{i\alpha}^{-1}$, and has a well-known interpretation in geometrical optics as a measure of focusing or defocusing: in particular, $|\det \Delta| \rightarrow \infty$ at a *caustic*, where light rays converge. Since $(N - N_0)$ is nonzero for $N \neq N_0$, a singularity in the van Vleck determinant implies a singularity in $\det \Gamma^{-1}$. Applying (12), we conclude

$$\frac{1}{\det \Gamma^{-1}} = \det \Gamma \equiv \Theta(N, N_0) = \exp \int_{N_0}^N \theta(N') dN'. \quad (19)$$

Therefore $\Theta \rightarrow 0$ at a caustic. This happens after finitely many e-folds only if $\theta \rightarrow -\infty$ during the flow. Otherwise, Θ is decreasing in regions where θ is negative, with large negative θ implying strong focusing. Large positive θ implies strong defocusing. More generally the propagator matrix can be rewritten in terms of Θ , giving

$$\Gamma_{ai} = \Theta(N, N_0)^{1/M} \mathcal{P} \exp \left(\int_{N_0}^N (\sigma + \omega)_{\alpha\beta}(N') dN' \right) \delta_{\beta i}. \quad (20)$$

The ordered exponential has determinant unity and therefore does not change the cross-sectional area of the bundle.

D. Adiabatic limit

Caustics have an important interpretation in the flows describing an inflationary model. If the bundle of trajectories has finite cross section, then the ensemble contains members which are evolving along multiple phase space trajectories. These are the eponymous “separate universes” with their individual initial conditions.

Under these circumstances one or more isocurvature modes exist. These are connecting vectors which relate the different ϕ_α within the bundle which all lie on a surface of fixed energy density, say Σ_ρ . Their number is determined by the rank of $h_{\alpha\beta} \Gamma_{\beta i}$. In the special case where the bundle cross-section decays to a point, there is a unique intersection between the bundle and Σ_ρ . Therefore $h_{\alpha\beta} \Gamma_{\beta i}$ has rank zero and all isocurvature modes disappear. In this limit, each member of the ensemble traverses the same orbit, differing from the others only by its relative position, which corresponds to the adiabatic mode, ζ . It follows that, when the cross-section collapses to a point, the fluctuations become purely adiabatic. Elliston et al. [18] described this as an ‘adiabatic limit’. After this limit has been reached ζ is conserved [34, 35].

Flows which reach an adiabatic limit during inflation are no more or less likely—or natural—from the viewpoint of fundamental physics. But flows reaching an adiabatic limit *are* more predictive, because a perturbation in the purely adiabatic mode remains adiabatic long after inflation ends [36], even during epochs for which we are ignorant of the relevant physics. Contrariwise, if any isocurvature modes remain then members of the ensemble may rearrange their relative positions until these modes decay. This possibility was emphasized by Meyers & Sivanandam [37]; see also

Ref. [18]. If the flow does not reach an adiabatic limit during inflation then the model is not predictive until we supply a prescription for the post-inflationary era, and observational predictions can depend on this choice.

Trivial, adiabatic and nonadiabatic caustics.—The outcome of this discussion is that approach to an adiabatic limit can be associated with convergence to a caustic. An early discussion of this principle, phrased almost precisely in these terms, was given by Wands & García-Bellido [21]. We conclude that $\Theta \rightarrow 0$ is a necessary condition for an adiabatic limit to occur, but as we now explain it is not sufficient.⁴ A caustic can be classified by the number of dimensions lost by the flow, or equivalently the number of null eigenvalues of the propagator Γ_{ai} at the caustic. An adiabatic limit is the special case where Γ_{ai} retains a *single* non-null eigenvalue, but $h_{\alpha\beta} \Gamma_{\beta i}$ has *no* non-null eigenvalues. We describe caustics which satisfy this condition as *adiabatic*.

Eq. (20) shows that, were the integrated shear and twist to remain bounded while $\Theta \rightarrow 0$, then $\Gamma_{ai} \rightarrow 0$. In this case no perturbations would survive, and we describe the caustic as *trivial*. An example is the case where $u_{\alpha\beta}$ is pure dilation. But barring an accurate cancellation of this kind, at least some component of $(\sigma + \omega)_{\alpha\beta}$ will typically scale proportionally to θ on approach to the caustic.⁵

Shear opposes focusing.—If the perturbations are not to vanish completely, then some *anisotropic* effect of shear and twist must oppose the isotropic contraction due to $\Theta \rightarrow 0$.

First suppose the twist is negligible. We assume that the eigenvectors of σ stabilize in the vicinity of the caustic. If the shear has some number of positive eigenvalues λ_i for which λ_i/θ has a finite, nonzero limit, then perturbations may survive in the subspace spanned by their corresponding eigenvalues. Tracelessness of σ implies that at least one eigenvalue must be negative, and perturbations in the subspace spanned by the corresponding eigenvectors will disappear. Hence, at least one dimension will be lost by the flow. In practice it is often simpler to work directly with the eigenvalues of the expansion tensor $u_{\alpha\beta}$.

If more than one eigenvalue of σ is positive, then perturbations may survive in a two- or higher dimensional subspace. In this case the caustic does not describe approach to an adiabatic limit, and we call it *nonadiabatic*. To obtain predictions for observable quantities the evolution must be

⁴ One may have some reservations about this conclusion, because it seems to violate the Liouville theorem which guarantees conservation of phase-space volume. However, it should be remembered that the *canonical* phase space coordinate to which Liouville’s theorem applies are not the field-space position and momenta which we are using. In particular, the canonical momenta will typically include powers of the scale factor a .

⁵ In principle $u_{\alpha\beta}$ could contain off-diagonal terms which grow *faster* than the diagonal terms, and therefore θ . In this case there could be a subspace of growing perturbations. If the growth is exponential this usually signals an instability, and the formalism we are describing becomes invalid.

continued. In practice this would require introduction of a reduced phase space describing only the surviving perturbations. The flow can then be followed in this reduced phase space until a further focusing event occurs. This may itself be an adiabatic limit, or might simply describe further reduction in the phase space. One should continue in this way until an adiabatic limit is finally achieved. An example of this behaviour could occur soon after the onset of slow-roll inflation. In the early stages, independent fluctuations in the field velocities survive. But when slow-roll is a good approximation these will be exponentially suppressed, making Θ become very small. One should therefore replace the full description by a reduced phase space which includes only field perturbations. In doing so one arrives at the field-space description of slow-roll inflation given in §II A.

Twist opposes focusing.—In slow-roll inflation, which we discuss in §II E below, a diverging shear is the only mechanism by which perturbations can survive on approach to a caustic. Where the twist is non-zero, which occurs when we do not impose the slow-roll approximation, more possibilities exist. Ultimately these must be addressed, to describe approach to an adiabatic limit when slow-roll is no longer a good approximation, but we defer this discussion for future work.

E. Focusing in the slow-roll approximation

In this subsection we give a more detailed discussion of the approach to a caustic during an era of slow-roll inflation.

Raychaudhuri equations.—Parametrizing each trajectory by e-folding number N , Eq. (4) constitutes an autonomous dynamical system. Therefore a derivative along the flow can be written $d/dN = u_\alpha \partial_\alpha$. In the absence of a nontrivial field-space metric all derivatives commute, and therefore $[\partial_\gamma, \partial_\beta]u_\alpha = 0$. Contracting with u_γ and rearranging terms, one finds

$$\frac{du_{\alpha\beta}}{dN} = \partial_\beta a_\alpha - u_{\alpha\gamma} u_{\gamma\beta}, \quad (21)$$

where a_α is the acceleration vector, defined by $a_\alpha = du_\alpha/dN = u_\beta \partial_\beta u_\alpha$. For a potential flow, this can be simplified; comparison with Eq. (4) shows that

$$a_\alpha = \frac{M_p^2}{2} \partial_\alpha v^2, \quad (22)$$

where, as above, v is the local refractive index.

The evolution equations for the dilation and shear can be written

$$\frac{d\theta}{dN} = M_p^2 \mathcal{H} - \frac{\theta^2}{M} - 2\sigma^2 \quad (23a)$$

$$\frac{d\sigma_{\alpha\beta}}{dN} = M_p^2 \left(\mathcal{H}_{\alpha\beta} - \frac{\mathcal{H}}{M} \delta_{\alpha\beta} \right) - \frac{2\theta}{M} \sigma_{\alpha\beta} - \left(\sigma_{\alpha\gamma} \sigma_{\gamma\beta} - \frac{2\sigma^2}{M} \delta_{\alpha\beta} \right). \quad (23b)$$

These are commonly known as the Raychaudhuri equations. An equation for the evolution of the twist could be found in the same way, but is not needed in the slow-roll approximation.

We have defined $\mathcal{H}_{\alpha\beta}$ to be the Hessian of v^2 ,

$$\mathcal{H}_{\alpha\beta} \equiv \frac{1}{2} \partial_\alpha \partial_\beta v^2, \quad (24)$$

and \mathcal{H} is its trace. Because the Hessian measures the local curvature of a function, one can regard $\mathcal{H}_{\alpha\beta}$ as a measure of the curvature of surfaces of constant refractive index in field space.

Jacobi equation.—Eq. (11) shows that Jacobi fields oriented along eigenvectors of $u_{\alpha\beta}$ with positive eigenvalues grow, whereas those oriented along eigenvectors with negative eigenvalues decay.

We can find an alternative description in terms of the refractive index v . Taking a derivative of (11) along the flow and using the Raychaudhuri equations to eliminate deriva-

tives of the dilation and shear yields the Jacobi equation,⁶

$$\frac{d^2 \delta \phi_\alpha}{dN^2} = M_p^2 \mathcal{H}_{\alpha\beta} \delta \phi_\beta. \quad (25)$$

It follows that the behaviour of the Jacobi fields is determined by the curvature of v^2 , considered as a function in field space. (Note this is related to, but not the same as, the curvature of surfaces of constant v .) Qualitatively, Jacobi fields oriented along eigenvectors of $\mathcal{H}_{\alpha\beta}$ with negative eigenvalues—directions of negative curvature—will have quasi-trigonometric solutions. These will pass through zero, corresponding to the collapse of some Jacobi fields to zero length. Fields oriented along eigenvectors with positive eigenvalues will have exponential solutions.

⁶ When using Jacobi fields to study geodesic deviation on a Riemannian manifold, this equation takes the form $\delta \ddot{\phi}_\alpha = -R_{\alpha\hat{n}\beta\hat{n}} \delta \phi_\beta$, where $R_{\alpha\hat{n}\beta\hat{n}} = \hat{n}^\rho \hat{n}^\sigma R_{\alpha\rho\beta\sigma}$ is a component of the Riemann curvature projected along the tangent to the geodesic.

Unless the initial conditions are precisely adjusted, these will typically grow.

Focusing theorem.—By adapting the geodesic focusing theorem of general relativity [23] we can determine the circumstances under which focusing will occur after finitely many e-folds. Pick a point on the flow where the expansion is negative, with value $\theta_* < 0$. Inspection of (23a) shows that, if $\mathcal{H} < 0$, then $\theta \rightarrow -\infty$ within $\Delta N = M/|\theta_*|$ e-folds, where M is the dimension of field space. Any point where $\theta = -\infty$ is a caustic, because on arrival at this point $\Theta = 0$.

Since Morse’s lemma implies that \mathcal{H} is negative in a neighbourhood of any local maximum of the refractive index, $v^2 = 2\epsilon$, one might hope to associate such local maxima with terminal points for inflation at which an adiabatic limit would be nearly achieved.

However, the conditions of the focusing theorem are not satisfied for typical potentials. More usually the slow-roll approximation forces all fields to settle into a terminal vacuum increasingly slowly, requiring an infinite number of e-folds to reach $\Theta = 0$. Moreover, in practical examples the slow-roll approximation will break down and inflation will terminate long before the caustic is reached. Therefore we should *not* expect to achieve precisely $\Theta = 0$ during inflation. Nevertheless, a model may be sufficiently predictive if the flow spends enough e-folds in a region of large negative θ that Θ is exponentially suppressed before inflation ends.

In simple potentials it is often clear when ζ ceases to evolve. But for more complicated potentials the situation may not be so clear. Within the slow-roll approximation, this discussion shows that $\Theta \gtrsim 1$ can be taken as a clear indication that isocurvature modes are still present. Their future decay is likely to influence ζ and the outcome of any calculation which terminates with $\Theta \gtrsim 1$ should not be considered a prediction for observable quantities. Conversely, $\Theta \ll 1$ is an indication that some decay of isocurvature modes has taken place. The precise nature of the decay must be deduced from the behaviour of the shear and twist. If perturbations survive only in a one-dimensional subspace then we can infer that the isocurvature modes have decayed to the point that ζ will be approximately conserved.

Example: quadratic Nflation.—We illustrate these ideas using the quadratic approximation to Nflation [38, 39]. The potential is

$$V = \sum_{\alpha} \frac{1}{2} m_{\alpha}^2 \phi_{\alpha}^2. \quad (26)$$

This model is of interest in its own right, but also describes the approach to a generic stable minimum after suitable choice of field space coordinates. We suppose that there is at least a modest hierarchy among the masses, and order these so that $m_{\alpha} < m_{\beta}$ if $\alpha < \beta$. The most massive field will settle into its minimum first, followed by the next most massive field. Therefore approach to the final minimum will be described by a trajectory on which only ϕ_1 is dynamical, with all other ϕ_{α} approximately zero. We describe this as the “inflow” trajectory.

On the inflow trajectory, the dilation satisfies

$$\theta^{\text{inf}} \approx -2 \frac{M_{\text{p}}^2}{\phi_1^2} \left(\sum_{\alpha \geq 2} \frac{m_{\alpha}^2}{m_1^2} - 1 \right). \quad (27)$$

The minimum $\phi_1 = 0$ is a caustic, but as discussed above it cannot be reached after finitely many e-folds (within the slow-roll approximation). The expansion tensor satisfies

$$u_{\alpha\beta}^{\text{inf}} \approx \begin{pmatrix} 2 \frac{M_{\text{p}}^2}{\phi_1^2} & & & \\ & \ddots & & \\ & & -2 \frac{m_{\alpha}^2}{m_1^2} \frac{M_{\text{p}}^2}{\phi_1^2} & \\ & & & \ddots \end{pmatrix}. \quad (28)$$

This has one positive eigenvalue and the rest negative, so we expect it will correspond to an adiabatic limit.

The (1,1) component of Γ^{inf} diverges near the caustic. This does not signal an instability, but only that $\delta\phi_1$ grows at precisely the required rate to give constant $\zeta \sim (H/\dot{\phi}_1)\delta\phi_1$.

Ordered exponentials such as (15) satisfy a composition property, allowing the integral over the inflationary trajectory to be broken in two. (See Fig. 2.) The first component is an integral from the initial point until the onset of the inflow trajectory. We take this to occur at $\phi_1 = \phi_1^*$, and choose ϕ_1^* so that (28) is a good approximation there. The propagator at this point is Γ_{ai}^* . It is a complicated weighted average over the trajectory, and cannot usually be calculated analytically. The second component is an integral over the inflow trajectory, which we denote $\Gamma_{\alpha\beta}^{\text{inf}}$. Therefore $\Gamma_{ai} = (\Gamma^{\text{inf}} \Gamma^*)_{ai}$. The inflow part can be computed from (28),

$$\Gamma^{\text{inf}} \approx \begin{pmatrix} \frac{\phi_1^*}{\phi_1} & & & \\ & \ddots & & \\ & & \left(\frac{\phi_1}{\phi_1^*} \right)^{m_{\alpha}^2/m_1^2} & \\ & & & \ddots \end{pmatrix}. \quad (29)$$

Except perhaps for special choices of initial conditions, Eq. (29) gives rank $r = 1$ at the caustic. Therefore this is an example of an adiabatic caustic.

In more general circumstances, it may be necessary to diagonalize $u_{\alpha\beta}^{\text{inf}}$ before integrating over the inflow trajectory. This is reminiscent of the introduction of scaling operators in a renormalization-group framework. Indeed, the entire analysis, and the emergence of rational but non-integer power-law scaling near the caustic, parallels a renormalization group analysis in the neighbourhood of a fixed point [40]; compare also Eq. (79) of Vernizzi & Wands [30].

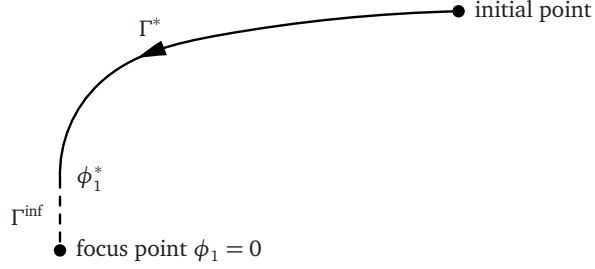


FIG. 2: Decomposition of propagator along an inflationary trajectory. Trajectories flowing into the minimum from most initial points join an “inflow trajectory” (represented by a dashed line) at $\phi_1 = \phi_1^*$. The precise location of the junction is initial-condition dependent. The inflow trajectory sinks into the caustic, which here is a focus point, giving nearly-universal behaviour in the final stages of approach. This parallels the discussion of universality in critical phenomena; however, here, the universal region is often physically inaccessible because the slow-roll approximation breaks down in the vicinity of the focus point. The remaining part of the trajectory (represented by a solid line) is non-universal, and typically cannot be calculated analytically.

Focusing in double quadratic model.—Away from the inflow trajectory it is usually necessary to proceed numerically. In Fig. 3 we show the evolution of the focusing parameters in the well-studied model of double quadratic inflation [20–22, 29, 30, 41, 42]. The potential is $V = m_1^2 \phi_1^2/2 + m_2^2 \phi_2^2/2$. We choose the mass ratio $m_1/m_2 = 9$ and set initial conditions $\phi_1 = 8.2M_P$ and $\phi_2 = 12.9M_P$.

Initially the evolution is mostly in the ϕ_2 direction. When ϕ_2 reaches the vicinity of its minimum there is a turn in field-space, which generates a spike in f_{NL} . After the turn, the inflow trajectory is reached along the ϕ_1 direction.

This evolution is reflected in the evolution of the bundle. Initially $\theta > 0$ and the cross-section slowly dilates. It reaches a maximum at roughly three times the original cross-sectional area. After the turn, θ rapidly drops to a negative value, and thereafter diverges exponentially to $-\infty$. Therefore the bundle-cross section very rapidly diminishes to almost zero cross-sectional area. This corresponds to an approximate caustic, and leads to an adiabatic limit.

Eventually the divergence in θ would be cut off by a breakdown of the slow-roll approximation, but for typical parameter choices Θ will already be exponentially small at this point.

Example: axion-quadratic model.—Elliston et al. [18] introduced an approximation to the hilltop region of axion N-flation [39]. The Hubble rate is dominantly supported by many axions in the quadratic region of their potential, and can be approximated by a single field. A few axions remain in the vicinity of the hilltop, where their contribution to H is negligible but their contribution to the three- and higher n -point functions in the adiabatic limit is large.

The potential is $V = m^2 \phi^2/2 + \Lambda^4(1 - \cos 2\pi\chi/f)$. We set $\Lambda^4 = 25m^2 f^2/4\pi^2$ and choose $f = M_P$. In Fig. 4 we show the evolution for initial conditions $\phi = 16M_P$ and $\chi = (f/2 - 0.001)M_P$.

The evolution is similar to the double quadratic model. Initially θ is positive and the cross-sectional area grows. At its peak, it is more than 200 times the original cross-section. Eventually ϕ approaches its minimum and the Hubble friction decreases to the point that χ can evolve. It rolls away from the hilltop, eventually ending inflation. During this phase θ switches sign, ultimately diverging exponentially to $-\infty$. Therefore we approach an adiabatic limit. However, Fig. 4c shows that the rate of approach is quite slow. The cross-section decays softly, and by the end of inflation $\Theta \sim 10^{-3}$. Therefore an approximate adiabatic limit is reached and we can expect the observables to be roughly conserved through the post-inflationary evolution.

III. TRANSPORT EQUATIONS

We now apply these ideas to obtain evolution (or “transport”) equations for the correlation functions in a fixed, comoving spacetime volume. In this section our analysis will be general, and can be applied to any perturbations whose evolution equations can be expressed in the form of Eq. (31). If necessary this can be achieved as described in §II by passing to a Hamiltonian framework. It follows that the transport of correlation functions is most naturally expressed in phase space.

Connecting vectors.—Consider the set-up described in §I, in which a comoving spacetime region of size μ is smoothed into separate universes of size L . Pick any one of these L -sized regions, which we take to be at spatial position \mathbf{x} . The separate universe approximation asserts that the evolution of the smoothed fields in this region is given by the flow equation (4). We denote the difference between these values and those in some other region, located at position $\mathbf{x} + \mathbf{r}$, by $\delta\phi_\alpha(\mathbf{r})$. This is a connecting vector in the sense of Eq. (11). Taylor expanding u_α , the corresponding deviation equation is

$$\frac{d\delta\phi_\alpha(\mathbf{r})}{dN} = u_{\alpha\beta}[\phi(\mathbf{x})]\delta\phi_\beta(\mathbf{r}) + \frac{1}{2}u_{\alpha\beta\gamma}[\phi(\mathbf{x})]\{\delta\phi_\beta(\mathbf{r})\delta\phi_\gamma(\mathbf{r}) - \langle\delta\phi_\beta(\mathbf{r})\delta\phi_\gamma(\mathbf{r})\rangle\} + \dots \quad (30)$$

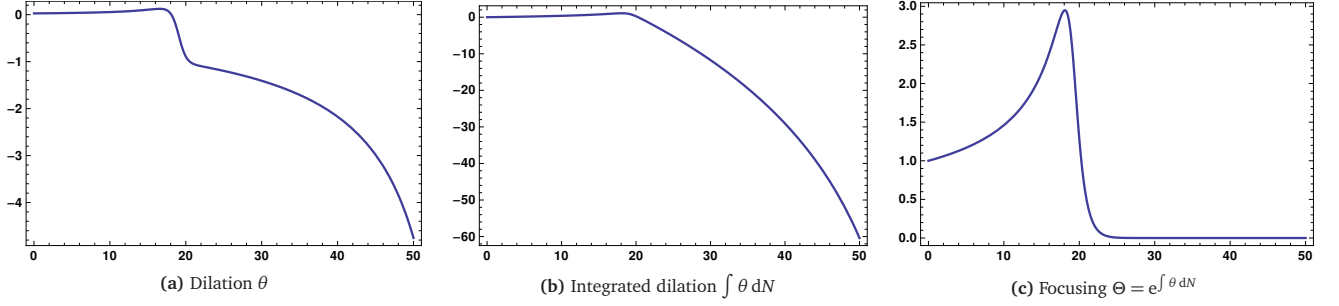


FIG. 3: Dilation, integrated dilation and focusing parameters in the double quadratic inflation model $V = \frac{1}{2}m_1^2\phi_1^2 + \frac{1}{2}m_2^2\phi_2^2$. The mass ratio is $m_1/m_2 = 9$, and the initial conditions are $\phi_1 = 8.2M_p$, $\phi_2 = 12.9M_p$. All plots are against the e-folding number N , measured from horizon exit of the mode in question.

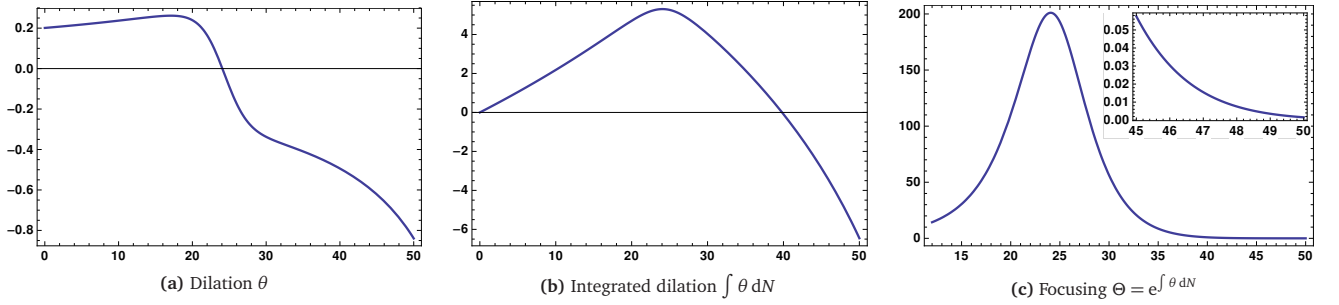


FIG. 4: Dilation, integrated dilation and focusing parameters in the axion-quadratic model $V = \frac{1}{2}m^2\phi^2 + \Lambda^4(1 - \cos 2\pi\chi/f)$. We have set $\Lambda^4 = 25m^2f^2/4\pi^2$ and $f = M_p$. The initial conditions are $\phi = 16M_p$, $\chi = (f/2 - 0.001)M_p$. In (c), the inset panel shows the evolution of Θ near the end of inflation. All plots are against the e-folding number N , measured from horizon exit of the mode in question.

We assume $\langle \delta\phi_\alpha(\mathbf{r}) \rangle = 0$ and have subtracted a zero-mode to preserve this throughout the motion.⁷ The tensor $u_{\alpha\beta}$ was defined in (11), and $u_{\alpha\beta\gamma} \equiv \partial_\gamma u_{\alpha\beta}$. We describe them, together with higher-index counterparts obtained by further differentiation, as u -tensors. They inherit a dependence on \mathbf{x} through evaluation at $\phi_\alpha = \phi_\alpha(\mathbf{x})$. After transformation to Fourier space, the subtractions in Eq. (30) correspond to discarding disconnected correlation functions. Therefore statistical properties of the ensemble do not depend on our choice of fiducial point.

If μ/L is not superexponentially large, we can typically expect $|\delta\phi_\alpha(\mathbf{r})|$ to be small and slowly varying. In Fourier space, this implies that $\delta\phi_\alpha(\mathbf{r})$ is constructed from only a few soft, infrared modes which we label \mathbf{k} . The remaining modes have been integrated out in the smoothing process used to obtain this effective, separate-universe description. Working explicitly in terms of these modes, Eq. (30) yields a connecting vector and deviation equation for each combination of species and \mathbf{k} -mode⁸

$$\frac{d\delta\phi_{\alpha'}}{dN} = u_{\alpha'\beta'}(\mathbf{x})\delta\phi_{\beta'} + \frac{1}{2}u_{\alpha'\beta'\gamma'}(\mathbf{x})\{\delta\phi_{\beta'}\delta\phi_{\gamma'} - \langle\delta\phi_{\beta'}\delta\phi_{\gamma'}\rangle\} + \dots \quad (31)$$

Eq. (31) has been written in an abbreviated “de Witt” notation, in which the primed, compound index α' carries *both* an unprimed species (or “flavour”) label α and a momentum \mathbf{k}_α . Contraction over primed indices implies summa-

tion over the flavour label and integration over the momentum label with measure d^3k . The 2- and 3-index u -tensors appearing here satisfy

$$u_{\alpha'\beta'}(\mathbf{x}) \equiv \delta(\mathbf{k}_\alpha - \mathbf{k}_\beta)u_{\alpha\beta}(\mathbf{x}) \quad (32a)$$

$$u_{\alpha'\beta'\gamma'}(\mathbf{x}) \equiv (2\pi)^{-3}\delta(\mathbf{k}_\alpha - \mathbf{k}_\beta - \mathbf{k}_\gamma)u_{\alpha\beta\gamma}(\mathbf{x}). \quad (32b)$$

Eq. (30) was given by Yokoyama et al. [7] in real space, and used to obtain evolution equations for the momentum-independent Lyth–Rodríguez Taylor coefficients. We explore the relationship between our approaches in Appendix A. However, Yokoyama et al. did not interpret

⁷ In the language of Feynman diagrams, this would correspond to removing contributions arising from disconnected pieces. This procedure is routine in applications of the separate universe principle.

⁸ In Eq. (31) we are keeping nonlinear terms in the evolution equation. We use the term “Jacobi field” to refer to infinitesimal connecting vectors, for which only the linear term need be kept.

$u_{\alpha\beta}$ as the expansion tensor of the flow or give the \mathbf{k} -space equations (31) and (32a)–(32b). As we will see, this \mathbf{k} -dependent information is necessary to obtain transport equations for the full set of coupled \mathbf{k} -space correlation functions.

One can arrive at the same conclusions using cosmological perturbation theory. Taking the background value of ϕ_α to be the average field over the μ -sized box, the perturbations within the box are $\delta\phi_\alpha(\mathbf{r})$. One should now interpret \mathbf{r} as a coordinate relative to the box. Restricting attention to the infrared modes in $\delta\phi_\alpha(\mathbf{r})$, for which k/a is negligible, we recover Eqs. (31) and (32a)–(32b).

Correlation functions.—The full set of connecting vectors contains all information required to determine evolution of the bundle, and therefore the evolution of all statistical quantities. In Eq. (31) this data is carried by the u -tensors. The transport equations for correlation functions are simply a reorganization of this information. Therefore they must also be expressible purely in terms of u -tensors. Since (31) shows that these tensors can be obtained by the separate universe argument or traditional perturbation theory, it follows that they will make equivalent predictions.

There are multiple ways to organize the u -tensors to produce evolution equations. In Ref. [8], transport equations were obtained after postulating a conservation equation for a probability density P ,

$$\frac{dP}{dN} + \partial_\alpha(u_\alpha P) = 0. \quad (33)$$

Evolution equations for the moments of P were extracted using both a Gauss–Hermite expansion, and generating functions. Here we describe a third, simpler method. Provided the perturbations can be treated classically, we expect $d\langle O \rangle/dN = \langle dO/dN \rangle$ for any quantity O .⁹

Two-point function.—We write the two-point function as $\Sigma_{\alpha'\beta'} \equiv \langle \delta\phi_{\alpha'} \delta\phi_{\beta'} \rangle$. Eq. (31) implies

$$\begin{aligned} \frac{d\Sigma_{\alpha'\beta'}}{dN} &= \left\langle \frac{d\delta\phi_{\alpha'}}{dN} \delta\phi_{\beta'} + \delta\phi_{\alpha'} \frac{d\delta\phi_{\beta'}}{dN} \right\rangle \\ &= u_{\alpha'\gamma'} \Sigma_{\gamma'\beta'} + u_{\beta'\gamma'} \Sigma_{\gamma'\alpha'} + [\geq 3 \text{ p.f.}] \\ &\equiv \{u, \Sigma\}_{\alpha'\beta'} + [\geq 3 \text{ p.f.}] \end{aligned} \quad (34)$$

where $\{A, B\}$ is the matrix anticommutator of A and B , and $[\geq 3 \text{ p.f.}]$ denotes terms including three-point functions or above which have been omitted. In general, the transport equations will couple correlation functions of all orders. They can be thought of as a limiting case of a Schwinger–Dyson hierarchy, applied to expectation values rather than the *in-out* amplitudes of scattering theory. Calzetta & Hu argued that the result could be interpreted as a Boltzmann hierarchy [43, 44].

As in any effective theory, the transport equations will be useful only if a reason can be found to systematically neglect an infinite number of terms. Applied to inflation, the statistical properties of the ensemble are nearly Gaussian: in the simplest models, an n -point function will typically be of order $H^{m(n)}$, where $m(n)$ is the smallest even integer at least as large as n [45]. This is suppressed compared to the natural scale M_p by $(H/M_p)^{m(n)} \ll 1$. However, this is not necessary; all that is required (or suggested by observation) for (34) to be valid is that the three- and higher n -point functions are substantially smaller than the two-point function.

Eq. (34) was given in Ref. [8] for an arbitrary n -field model, but with $\Sigma_{\alpha'\beta'}$ interpreted as the real-space correlation function. The single-field case is discussed by Gardiner [46]. With the u -tensors given in (32a)–(32b), Eq. (34) applies for the full \mathbf{k} -dependent correlation function.

Three-point function.—We write the three-point function as $\alpha_{\alpha'\beta'\gamma'} \equiv \langle \delta\phi_{\alpha'} \delta\phi_{\beta'} \delta\phi_{\gamma'} \rangle$. Keeping contributions of order $O(\Sigma^2)$ and $O(\alpha)$, we conclude

$$\begin{aligned} \frac{d\alpha_{\alpha'\beta'\gamma'}}{dN} &= u_{\alpha'\lambda'} \alpha_{\lambda'\beta'\gamma'} + u_{\alpha'\lambda'\mu'} \Sigma_{\lambda'\beta'} \Sigma_{\mu'\gamma'} \\ &\quad + \text{cyclic } (\alpha' \rightarrow \beta' \rightarrow \gamma') + [\geq 4 \text{ p.f.}] \end{aligned} \quad (35)$$

In simple models, the scaling estimate $\langle \delta\phi^n \rangle \sim H^{m(n)}$ makes both terms the same order of magnitude. For (35) to be valid requires the 4-point function to be substantially smaller, which is also supported by observation [47].

IV. EVOLUTION OF CORRELATION FUNCTIONS

A. Solution of the transport hierarchy by raytracing

The transport equations (34) and (35) can be solved using the machinery developed in §II. The key ingredients are the phase-space flows which describe evolution of individual “separate universes,” and the Jacobi fields which connect them. The solution is formal and depends only on the structure described in §III. Therefore there is no requirement to impose the slow-roll approximation, and when written over the full phase-space our equations apply quite generally. When truncated to field-space they reproduce the slow-roll evolution.

Two-point function.—We write the two-point function $\Sigma_{\alpha'\beta'}$ in the form

$$\Sigma_{\alpha'\beta'} \equiv \Gamma_{\alpha'i'} \Gamma_{\beta'j'} \Sigma_{i'j'}, \quad (36)$$

where Γ is to be determined. This notation has been chosen because Γ will turn out to be the propagator matrix (15) for the primed indices. Indeed, (36) is a solution of the

⁹ This equation both implies and is implied by conservation of probability, Eq. (33)

transport equation (34) if

$$\frac{d\Gamma_{\alpha'i'}}{dN} = u_{\alpha'\gamma'}\Gamma_{\gamma'i'} \quad (37a)$$

$$\frac{d\Sigma_{i'j'}}{dN} = O(H^4) \approx 0. \quad (37b)$$

Eq. (37a) is the equation for a Jacobi field, Eq. (11).

In writing (37b) we have assumed approximate Gaussianity, so that contributions from higher-order correlation functions are suppressed by at least a power of H^2 compared to the terms which have been retained. Keeping these terms would yield the “loop corrections” of the Lyth–Rodríguez formalism [19, 48, 49]. To the order we are working, $\Sigma_{i'j'}$ should be identified as a constant: it is the value of the two-point function evaluated at $N = N_0$, where N_0 is the initial time which appears in the propagator (15). We write this constant value $\mathcal{S}_{i'j'}$.

The primed propagator satisfies

$$\Gamma_{\alpha'i'} = \delta(\mathbf{k}_\alpha - \mathbf{k}_i)\Gamma_{ai}, \quad (38)$$

where Γ_{ai} is the flavour propagator (15). Therefore, written more explicitly, Eq. (36) becomes

$$\langle \delta\phi_\alpha(\mathbf{k}_\alpha)\delta\phi_\beta(\mathbf{k}_\beta) \rangle = \Gamma_{ai}\Gamma_{\beta j}\langle \delta\phi_i(\mathbf{k}_\alpha)\delta\phi_j(\mathbf{k}_\beta) \rangle_0, \quad (39)$$

where our usual convention—that Latin indices denote evaluation of the correlation function at some initial time N_0 —continues to apply. For the two point function, practical calculations usually simplify if this is taken to be the horizon-crossing time associated with scale $k = |\mathbf{k}_\alpha| = |\mathbf{k}_\beta|$. We have indicated this by attaching a subscript ‘0’ to the correlation function. With this understanding, and recollecting the identification (16), Eq. (39) is the familiar “ δN ” result [4, 5].

Three-point function.—Similar methods can be used to solve for the three- and four-point functions. We write $\alpha_{\alpha'\beta'\gamma'} \equiv \Gamma_{\alpha'i'}\Gamma_{\beta'j'}\Gamma_{\gamma'k'}\alpha_{i'j'k'}$. As for the two-point function, the propagator matrices absorb contributions from the $u_{\alpha\beta}$ -tensors. In the case of $\Sigma_{\alpha'\beta'}$ there were no other terms, making the “kernel” $\Sigma_{i'j'}$ time independent. Here, the presence of terms involving u 3-tensors provides a source for $\alpha_{i'j'k'}$. We find¹⁰

$$\frac{d\alpha_{i'j'k'}}{dN} = (\Gamma_{i'\alpha'}^{-1}u_{\alpha'\beta'\gamma'}\Gamma_{\beta'm'}\Gamma_{\gamma'n'})\mathcal{S}_{m'j'}\mathcal{S}_{n'k'} + \text{cyclic} + O(H^6), \quad (40)$$

where, as above, \mathcal{S}_{ij} is the initial value of the two-point function introduced in (36). The estimate $O(H^6)$ for the truncation error, beginning with contributions from the

four-point function, again assumes that the correlation functions order themselves in even powers of H . We define the matrix $\Gamma_{i'\alpha'}^{-1}$ to be the left-inverse of the propagator, $\Gamma_{i'\alpha'}^{-1}\Gamma_{\alpha'j'} = \delta(\mathbf{k}_i - \mathbf{k}_j)\delta_{ij}$. Inspection of (38) shows that it can be written

$$\Gamma_{i'\alpha'}^{-1} = \delta(\mathbf{k}_i - \mathbf{k}_\alpha)\Gamma_{i\alpha}^{-1}, \quad (41)$$

where $\Gamma_{i\alpha}^{-1}$ is the conventional matrix inverse of the flavour propagator, Eq. (15). In what follows it is useful to define a projected u 3-tensor, $\tilde{u}_{i'j'k'}$, by

$$\tilde{u}_{i'j'k'} = \Gamma_{i'\alpha'}^{-1}u_{\alpha'\beta'\gamma'}\Gamma_{\beta'j'}\Gamma_{\gamma'k'}. \quad (42)$$

Combining (38) and (41), it follows that the explicit \mathbf{k} - and flavour-dependence can be written

$$\tilde{u}_{i'j'k'} = \delta(\mathbf{k}_i - \mathbf{k}_j - \mathbf{k}_k)\tilde{u}_{ijk}, \quad (43)$$

where the tensor \tilde{u}_{ijk} is the obvious flavour projection of u_{ijk} , so that $\tilde{u}_{ijk} = \Gamma_{i\alpha}^{-1}u_{\alpha\beta\gamma}\Gamma_{\beta j}\Gamma_{\gamma k}$.

With these definitions, Eq. (40) can be solved by quadrature. Up to loop corrections, we find

$$\alpha_{i'j'k'} = \mathcal{A}_{i'j'k'} + \int_{N_0}^N \tilde{u}_{i'm'n'}(N')\mathcal{S}_{m'j'}\mathcal{S}_{n'k'} dN' + \text{cyclic}, \quad (44)$$

where $\mathcal{A}_{i'j'k'}$ should be regarded as the value of the three-point function at $N = N_0$. The complete solution can be written (again up to loop corrections)

$$\alpha_{\alpha'\beta'\gamma'} = \Gamma_{\alpha'i'}\Gamma_{\beta'j'}\Gamma_{\gamma'k'}\mathcal{A}_{i'j'k'} + \left(\Gamma_{\alpha'm'n'}\Gamma_{\beta'j'}\Gamma_{\gamma'k'}\mathcal{S}_{m'j'}\mathcal{S}_{n'k'} + \text{cyclic} \right), \quad (45)$$

where the cyclic permutations exchange $\alpha' \rightarrow \beta' \rightarrow \gamma'$.

One can regard Eqs. (37a)–(37b) and (44) as analogous to the “line of sight” integral which is used to obtain a formal solution to the Boltzmann equation in calculations of the cosmic microwave background anisotropies.

The quantity $\Gamma_{\alpha'm'n'}$ is defined by

$$\Gamma_{\alpha'm'n'} \equiv \Gamma_{\alpha'i'} \int_{N_0}^N \tilde{u}_{i'm'n'}(N') dN'. \quad (46)$$

Observe that Eq. (46) is symmetric in the indices m' and n' . With our choices for the \mathbf{k} - and flavour-dependence of its constituent quantities, it can be written

$$\Gamma_{\alpha'm'n'} = \delta(\mathbf{k}_\alpha - \mathbf{k}_m - \mathbf{k}_n)\Gamma_{amn}, \quad (47)$$

where Γ_{amn} is the flavour-only object obtained by exchanging primed for unprimed indices in (46). Comparing with (15), it follows that (up to matrix ordering ambiguities) Γ_{amn} is the derivative of the propagator,

$$\frac{\partial^2 \phi_\alpha}{\partial \phi_m \partial \phi_n} = \Gamma_{amn}. \quad (48)$$

Eq. (45) can now be recognized as the Lyth–Rodríguez formula for the three-point function [5].

¹⁰ We are allowing $\alpha_{i'j'k'}$ to be a function of N , which means our index convention must be interpreted more abstractly. The expressions for Γ -matrices to which Eq. (40) leads, such as Eqs. (49a)–(49b), can be interpreted in the original sense.

B. Flow equations for “ δN ” coefficients

We conclude that the transport equations (34) and (35) are equivalent to the Taylor expansion algorithm of Lyth & Rodríguez for the three-point function. Also, because the u -tensors could equally well be derived using the methods of cosmological perturbation theory, all these methods will give answers which agree. Within this narrow reading, our analysis can be interpreted as a demonstration that these methods are interchangeable. Therefore we believe that statements to the effect that any particular method currently in use has an intrinsic drawback when compared with another, *as a matter of principle*, are wrong.

Nevertheless it is true that some approaches have advantages in practice, although no one approach outperforms the others in all applications. For example, as explained in §I, in some models the Taylor expansion algorithm leads to very simple analytic formulae. This property has encouraged a large literature studying models to which the method can be applied.

In this broader context our analysis is not simply a reformulation of existing results. First, as a byproduct of the raytracing method we have obtained explicit (but formal) expressions for the Lyth–Rodríguez Taylor coefficients,

$$\frac{\partial \phi_\alpha}{\partial \phi_i} = \Gamma_{\alpha i} = \mathcal{P} \exp \left(\int_{N_0}^N u_{\alpha\beta}(N') dN' \right) \delta_{\beta i} \quad (49a)$$

$$\frac{\partial^2 \phi_\alpha}{\partial \phi_i \partial \phi_j} = \Gamma_{\alpha ij} = \Gamma_{\alpha m} \int_{N_0}^N \tilde{u}_{mij}(N') dN'. \quad (49b)$$

Analytically, the Taylor expansion method is useful only when a solution to (49a) can be found in closed form. This has been achieved only for a limited class of potentials obeying some form of separability criteria; a summary appears in Ref. [18] together with references to the original literature. Eq. (49a) clarifies the difficulty encountered in obtaining analytic formulae as the difficulty of computing closed-form expressions for a path-ordered exponential. A sophisticated theory is available [50] but explicit expressions can usually be obtained only in special cases, or where the expansion tensor commutes with itself at different times. It is possible that Eq. (49a) could be used to extend analytic progress beyond the separable cases, but we have not investigated this possibility in detail.

Eqs. (49a)–(49b) were given, in slightly different notation, by Yokoyama et al. [7]. Because of its close relation to the present discussion we review and extend the Yokoyama et al. approach in Appendix A.

Second, a naïve numerical implementation of the Taylor expansion formula is unfavourable. Beginning with fractionally displaced initial conditions one must evolve the equations of motion over many e-folds, during which numerical noise is accumulating. Taking differences between these evolved solutions requires high-accuracy integration in order that the small displacement in initial conditions is not swamped by noise. The explicit solutions (49a)–(49b) allow this naïve approach to be replaced by a simple sys-

tem of ordinary differential equations for $\Gamma_{\alpha i}$ and $\Gamma_{\alpha ij}$. The $\Gamma_{\alpha i}$ equation is the Jacobi equation (37a), after dropping primes on indices. The initial condition is $\Gamma_{\alpha i} = \delta_{\alpha i}$. The $\Gamma_{\alpha ij}$ equation can be obtained by differentiation of (49b). It is

$$\frac{d\Gamma_{\alpha ij}}{dN} = u_{\alpha\beta} \Gamma_{\beta ij} + u_{\alpha\gamma} \Gamma_{\beta i} \Gamma_{\gamma j}, \quad (50)$$

with initial condition $\Gamma_{\alpha ij} = 0$.

The same approach can be applied systematically to deduce transport equations for any of the Taylor coefficients. Yokoyama et al. wrote the transport equation (37a) for $\Gamma_{\alpha i}$, but did not write (50) for $\Gamma_{\alpha ij}$ which they computed directly from (49b). See Appendix A for a comparison.

C. Transport of “shape” amplitudes

The results of §IV A apply for arbitrary initial conditions $\mathcal{S}_{i'j'}$, $\mathcal{A}_{i'j'k'}$ for the two- and three-point functions. But for application to inflation, we will usually wish to apply them to the correlation functions produced in a specific model. In this case the fields ϕ_α will be a collection of light scalars for which $\mathcal{S}_{i'j'}$ and $\mathcal{A}_{i'j'k'}$ can be computed using the *in-in* formulation of quantum field theory [26]. These yield very specific k -dependences whose amplitudes we wish to track.

In this section, our analysis remains general and continues to apply to the full phase space.

Two-point function.—The two-point function is straightforward. For a nearly scale-invariant spectrum we have

$$\Sigma_{\alpha'\beta'} \equiv (2\pi)^3 \delta(\mathbf{k}_\alpha + \mathbf{k}_\beta) \frac{\Sigma_{\alpha\beta}}{k^3}, \quad (51)$$

where $k = |\mathbf{k}_\alpha| = |\mathbf{k}_\beta|$ and the flavour matrix $\Sigma_{\alpha\beta}$ should be nearly independent of k . Transport of $\Sigma_{\alpha\beta}$ can be accomplished using (39), or simply by solving the transport equation (34) with an appropriate initial condition after dropping primes on indices. That gives

$$\Sigma_{\alpha\beta} = \Gamma_{\alpha i} \Gamma_{\beta j} \mathcal{S}_{ij}, \quad (52)$$

where \mathcal{S}_{ij} is the initial value of $\Sigma_{\alpha\beta}$. The mild k -dependence of (52) can also be obtained using transport techniques [51].

Three-point function.—Here, more possibilities exist. It is known that the $\mathcal{O}(\mathcal{S}^2)$ terms in (45) dominate whenever the bispectrum is large enough to be observed [30, 52]. Eq. (45) shows that these contributions add incoherently to the contribution from \mathcal{A}_{ijk} , so they can be studied separately. Using (51) and overall symmetry of the correlation function under exchange of indices, we can write

$$\alpha_{\alpha'\beta'\gamma'} \supseteq (2\pi)^3 \delta(\mathbf{k}_\alpha + \mathbf{k}_\beta + \mathbf{k}_\gamma) \left(\frac{\alpha_{\alpha|\beta\gamma}}{k_\beta^3 k_\gamma^3} + \frac{\alpha_{\beta|\alpha\gamma}}{k_\alpha^3 k_\gamma^3} + \frac{\alpha_{\gamma|\alpha\beta}}{k_\alpha^3 k_\beta^3} \right), \quad (53)$$

where the notation “ \supseteq ” indicates that the three-point contribution contains this contribution among others. The amplitudes $a_{\alpha|\beta\gamma}$ are symmetric under exchange of β and γ , but not otherwise. Using Eqs. (32a), (32b), (35) and (51), we find the transport equation

$$\frac{d\alpha_{\alpha|\beta\gamma}}{dN} = u_{\alpha\lambda}\alpha_{\lambda|\beta\gamma} + u_{\beta\lambda}\alpha_{\alpha|\lambda\gamma} + u_{\gamma\lambda}\alpha_{\alpha|\beta\lambda} + u_{\alpha\lambda\mu}\Sigma_{\lambda\beta}\Sigma_{\mu\gamma}. \quad (54)$$

If desired, we can apply the same method of formal solution described in §IVA. This yields

$$\alpha_{\alpha|\beta\gamma} = \Gamma_{\alpha mn}\Gamma_{\beta j}\Gamma_{\gamma k}\mathcal{S}_{mj}\mathcal{S}_{nk}. \quad (55)$$

In combination with (53) this reproduces our earlier for-

mula (45), neglecting the initial contribution $\mathcal{A}_{i'j'k'}$.

D. Connections between the transport and other approaches

Up to this point we have shown that the Jacobi fields which connect “separate universe” trajectories in phase space can be used to solve the transport equations for the full set of \mathbf{k} -space correlation functions. But as we have explained, the transport hierarchy is just one of many techniques for handling correlation functions. We now pause to examine the connections between these approaches.

δN formalism.—In the Lyth–Rodríguez approach, or “ δN formalism”, one makes a Taylor expansion of the field values on a final hypersurface in terms of field values on some initial hypersurface. Following the discussion surrounding Eq. (30), and with the same meaning for the vectors \mathbf{x} and \mathbf{r} , this can be written

$$\delta\phi_{\alpha}(\mathbf{r}) = \Gamma_{\alpha i}(\mathbf{x})\delta\phi_i(\mathbf{r}) + \frac{1}{2}\Gamma_{\alpha ij}(\mathbf{x})\{\delta\phi_i(\mathbf{r})\delta\phi_j(\mathbf{r}) - \langle\delta\phi_i(\mathbf{r})\delta\phi_j(\mathbf{r})\rangle\} + \dots \quad (56)$$

Note that, despite appearances, we are making no assumption that the evolution of $\delta\phi$ is close to an attractor. Therefore there is no requirement to invoke the slow-roll approximation. It is true that the existence of an attractor would make the canonical momenta purely a function of the fields, yielding an equation with the appearance of Eq. (56). But as we have explained, by working in a first-order Hamiltonian formalism we can obtain expressions such as (56) without this limitation. Therefore we allow the $\delta\phi_i$ to include perturbations of the canonical momenta if necessary, in which case the indices α, i , etc. range over the $2M$ dimensions of phase space. Where slow-roll is a good approximation we can revert to a simpler formulation based on field space.

We have already remarked that the Γ -tensors are the derivatives (16) and (48). In Eq. (56) the $\delta\phi_{\alpha}$ are all defined on spatially flat hypersurfaces. More commonly, an analogous expansion is made for the total e-folding number N , measured from a flat slice to a final comoving slice; we give an explicit relation in §V. The choice of slicing simply corresponds to the gauge in which we wish to work [8].

For (56) to be useful, some means must be found to compute $\Gamma_{\alpha i}$ and $\Gamma_{\alpha ij}$.

Flow equations.—As a by-product of the raytracing solution, or “line of sight” integral, we obtained the evolution equations (37a) and (50). These allow the Γ -tensors to be computed easily. However, the same equations can be obtained directly from the separate universe formula, Eq. (56). Substituting (56) into both the right- and left-hand sides of (30) and separating the resulting expansion order-by-order, we immediately arrive at Eqs. (37a) and

(50). This still does not require the slow-roll approximation.

Transfer matrices.—We have observed that Eq. (30) arises in the $k/aH \rightarrow 0$ limit of cosmological perturbation theory (“CPT”). Within that framework, at least in the first-order theory, it is common to introduce “transfer matrices” which relate field perturbations at different times [9]. Typically these are chosen to be the adiabatic and isocurvature directions, but in principle any basis can be used.

Restricting to first-order, the transfer matrix is determined precisely by the leading term of (56), or a gauge transformation of it. It follows that Eq. (56) represents the extension of the transfer matrix to second-order (and beyond), and Eqs. (37a) and (50) give the evolution of the transfer tensors $\Gamma_{\alpha i}$, $\Gamma_{\alpha ij}$. Therefore the transfer-matrix formalism is precisely equivalent to the separate universe picture and traditional cosmological perturbation theory. Note that if the perturbations are projected onto adiabatic and isocurvature modes this requires use of the correct u tensors at each time step.

CPT implies transport equations.—Finally, we show that cosmological perturbation theory implies the transport hierarchy with which we began. We write

$$\Sigma_{\alpha\beta} = \Gamma_{\alpha i}\Gamma_{\beta j}\mathcal{S}_{ij}, \quad (57)$$

which, neglecting “loops,” follows from (56) and therefore either CPT or a transfer-matrix approach. Differentiating both sides with respect to time, recalling that \mathcal{S}_{ij} is time-

independent, and make use of (37a) we find

$$\frac{d\Sigma_{\alpha\beta}}{dN} = (u_{\alpha\mu}\Gamma_{\mu i}\Gamma_{\beta j} + u_{\beta\mu}\Gamma_{\mu i}\Gamma_{\alpha j}) \mathcal{S}_{ij}. \quad (58)$$

This gives the transport equation for $\Sigma_{\alpha\beta}$, Eq. (34). A similar procedure leads to the transport equation for $\alpha_{\alpha\beta\gamma}$, Eq. (34). It follows that each of these approaches implies and is implied by the others.

V. GAUGE TRANSFORMATIONS

To this point, the formalism we have developed enables the correlation functions of fluctuations in the fields and their momenta, $\delta\phi_\alpha$ and δp_α , to be evolved along the bundle of trajectories picked out by an ensemble of smoothed regions. However, by themselves these fluctuations are not observable. Only specific combinations are observable, of which the most important is the primordial curvature fluctuation ζ . Therefore to proceed we require expressions for the gauge transformation between the $\delta\phi_\alpha$, δp_α and ζ .

In this section we impose the slow-roll approximation throughout, enabling us to work on field space and make use of the hypersurface-orthogonal property of the flow. We intend to return to the general case in a future publication.

A. Explicit transformations

In the slow-roll approximation there is no need to track the momentum fluctuations δp_α , which are purely determined by the field fluctuations $\delta\phi_\alpha$. Therefore ζ can be written purely in terms of the field fluctuations.

On superhorizon scales, the appropriate gauge transformation can be written as a Taylor expansion,

$$\zeta = N_\alpha \delta\phi_\alpha + \frac{1}{2} N_{\alpha\beta} (\delta\phi_\alpha \delta\phi_\beta - \langle \delta\phi_\alpha \delta\phi_\beta \rangle) + \dots, \quad (59)$$

where all fields are evaluated at the same spatial position and a constant has been subtracted to set $\langle \zeta \rangle = 0$. The Taylor coefficients N_α and $N_{\alpha\beta}$ have been given by various authors [8, 53]. Working in field space, we give a purely geometrical derivation. This argument relies on the property that the flow is orthogonal to surfaces of constant density in field space, and therefore will not generalize directly to the full phase space.

Linear term.—Consider Fig. 5a. We wish to compute the coefficient N_α at a field-space point x , which can be taken to lie on a hypersurface of fixed energy density ρ . We denote this hypersurface Σ_ρ . According to the separate universe approximation, N_α can be computed from the number of e-folds required to flow back to Σ_ρ after making a generic (“off-shell”) displacement from x . Anticipating the discussion of second-order contributions, we denote this displacement $\delta\phi^1$ and write $z = x + \delta\phi^1$.

The number of e-folds required to return to Σ_ρ must be computed along the inflationary trajectory which passes

through z . In Fig. 5a, this trajectory intersects Σ_ρ at y . The tangent to the trajectory at y is the normal vector $\hat{n}(y)$. Therefore the (“on-shell”) field-space displacement along this trajectory, to first order in $\delta\phi^1$, is $\delta\phi_\alpha^{\text{flow}} \approx -\hat{n}_\alpha \hat{n}_\beta \delta\phi_\beta^1$. The symbol ‘ \approx ’ denotes equality up to higher-order terms in $\delta\phi^1$ which have been omitted, and we have adopted a convention in which quantities evaluated at x —such as the unit vector \hat{n} —are written without an argument. Combining Eqs. (3) and (5), we conclude

$$\delta N \approx -\frac{1}{M_p} \frac{\hat{n}_\alpha \delta\phi_\alpha^1}{\sqrt{2\epsilon}} \quad (60)$$

and therefore

$$\frac{\partial N}{\partial \phi_\alpha^1} = -\frac{1}{M_p} \frac{\hat{n}_\alpha}{\sqrt{2\epsilon}} = -\frac{1}{M_p} \frac{\hat{n}_\alpha}{v}. \quad (61)$$

where we have reintroduced the refractive index $v = \sqrt{2\epsilon}$ defined in §II A. Eq. (61) is the term N_α in (59).

Quadratic term.—The quadratic Taylor coefficient can be obtained from the variation in $\partial N / \partial \phi_\alpha^1$ under a *second* generic displacement $\delta\phi^2$. Under this displacement the origin is shifted to $x' = x + \delta\phi^2$. Because the energy density at x' will typically differ from ρ , it lies on a displaced hypersurface $\Sigma_{\rho'}$. However, the definition of N is unchanged and must still be measured to the intersection with Σ_ρ at y . We should compute the flow along the trajectory passing through z . The path $z \rightarrow y' \rightarrow y$ is a discrete approximation to an integral along this flow. The calculation should be carried to linear order in $\delta\phi^1$ and $\delta\phi^2$ independently.

In Fig. 5b, the on-shell flow from $z = x' + \delta\phi^1$ back to $\Sigma_{\rho'}$ is $\delta\phi^a$. Repeating the analysis above, we find

$$\delta\phi_\alpha^a \approx -\hat{n}'_\alpha \hat{n}'_\beta \delta\phi_\beta^1, \quad (62)$$

where $\hat{n}'_\alpha \equiv \hat{n}_\alpha(x') \approx \hat{n}_\alpha + \delta\phi_\beta^2 \partial_\beta \hat{n}_\alpha$. (It is only necessary to work to first order in $\delta\phi^2$, since (62) is proportional to $\delta\phi^1$.) The on-shell flow from y' back to $\Sigma_{\rho'}$ is

$$\delta\phi_\alpha^b \approx -\hat{n}''_\alpha \hat{n}''_\beta \Delta_\beta - \hat{n}''_\alpha \left(\frac{K_{\beta\gamma}}{2} - \partial_\beta \hat{n}_\gamma \right) \Delta_\beta \Delta_\gamma. \quad (63)$$

We have defined Δ_α to be the displacement to y' ,

$$\Delta_\alpha \equiv \delta\phi_\alpha^1 + \delta\phi_\alpha^2 + \delta\phi_\alpha^a, \quad (64)$$

and $\hat{n}''_\alpha \equiv \hat{n}_\alpha(y')$. The symmetric tensor $K_{\alpha\beta}$ is the extrinsic curvature of Σ_ρ , or “second fundamental form,” and is defined by $K_{\alpha\beta} \equiv h_{\alpha\gamma} h_{\beta\delta} \partial_\gamma \hat{n}_\delta$ [23]. It is related to the dilation and shear of the expansion tensor via

$$K_{\alpha\beta} = \frac{1}{M_p v} \left(\frac{\theta}{d} h_{\alpha\beta} + \sigma_{\alpha\beta}^{\text{iso}} \right), \quad (65)$$

where $\sigma_{\alpha\beta}^{\text{iso}}$ is the projection of the shear onto the isocurvature subspace, $\sigma_{\alpha\beta}^{\text{iso}} \equiv h_{\alpha\gamma} h_{\beta\delta} \sigma_{\gamma\delta}$. The first term in (63)

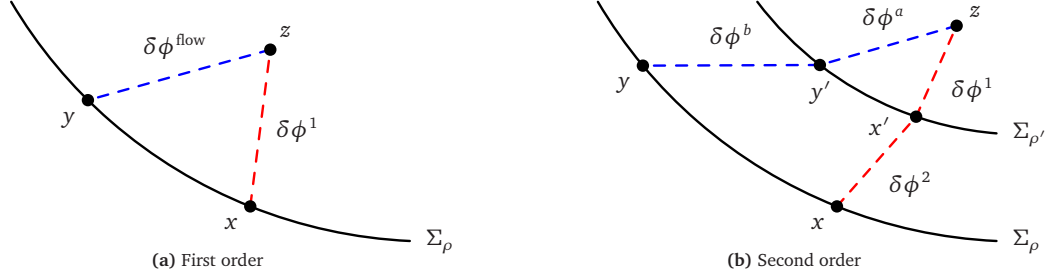


FIG. 5: Gauge transformations in field space

is a linear, trigonometric approximation. The second is a correction for the curvature of Σ_ρ . A similar construction

could be used to obtain the Taylor coefficients at any desired order.

After computing all appropriate variations, we find

$$\begin{aligned} \frac{\partial^2 N}{\partial \phi_\alpha^1 \partial \phi_\beta^2} &= -\frac{1}{M_P} \left(\frac{K_{\alpha\beta}}{\sqrt{2\epsilon}} + \hat{n}_\alpha \partial_\beta (2\epsilon)^{-1/2} + \hat{n}_\beta \partial_\alpha (2\epsilon)^{-1/2} - \hat{n}_\alpha \hat{n}_\beta \hat{n}_\gamma \partial_\gamma (2\epsilon)^{-1/2} \right) \\ &= -\frac{1}{M_P \nu} \left(K_{\alpha\beta} - \hat{n}_\alpha \mathcal{D}_\beta \ln \nu - \hat{n}_\beta \mathcal{D}_\alpha \ln \nu - \frac{\hat{n}_\alpha \hat{n}_\beta}{M_P} \frac{\eta}{2\nu} \right), \end{aligned} \quad (66)$$

where $\eta = d \ln \epsilon / dN$ is the natural generalization of the single-field η -parameter. It measures the variation of ϵ along the adiabatic direction. To yield sufficient e-foldings, it must typically be small while observable scales are leaving the horizon. Defining $\mathcal{D}_\sigma \equiv \hat{n}_\alpha \partial_\alpha$ to be a derivative along \hat{n}_α , it can be written

$$\eta \equiv \frac{2M_P}{\nu} \mathcal{D}_\sigma \epsilon. \quad (67)$$

In addition, $\mathcal{D}_\alpha \equiv h_{\alpha\beta} \partial_\beta$ is a derivative in the plane tangent to Σ_ρ at x . This tangent space can be interpreted as the subspace of isocurvature modes. Only the η -component of (66) depends purely on the local behaviour of the adiabatic direction, and therefore the direction in field space restricted by the slow-roll approximation. The remaining terms all probe details of the isocurvature subspace.

Dropping the distinction between $\delta\phi^1$ and $\delta\phi^2$, Eq. (66) is equal to $N_{\alpha\beta}$. It is symmetric even though we have not treated the displacements $\delta\phi^1$ and $\delta\phi^2$ equally. This is a consequence of associativity of vector addition, which makes z the same no matter in which order we apply the displacements. The inflationary trajectory passing through z is unique, so $N_{\alpha\beta}$ can only depend on a symmetric combination of $\delta\phi^1$ and $\delta\phi^2$.

Eq. (66) shows that $N_{\alpha\beta}$ depends on the anisotropy of ϵ —or, in the optical interpretation, the refractive index ν . It also depends on the extrinsic curvature of Σ_ρ , which is a function of the shape of the hypersurfaces of constant energy density. In particular, because $\hat{n}_\alpha K_{\alpha\beta} = 0$, this term

can be interpreted as a metric on the subspace of isocurvature modes.

B. Local mode f_{NL}

Two-point function.—These results can be combined to obtain the usual formulae for the amplitude of the local mode, f_{NL} . With our usual assumptions about the amplitude of those correlation functions we neglect, the two-point function of ζ satisfies

$$\begin{aligned} \langle \zeta(\mathbf{k}_1) \zeta(\mathbf{k}_2) \rangle &= (2\pi)^3 \delta(\mathbf{k}_1 + \mathbf{k}_2) N_\alpha N_\beta \Gamma_{\alpha i} \Gamma_{\beta j} \frac{\mathcal{S}_{ij}}{k^3} \\ &+ \mathcal{O}\left(\frac{H^4}{M_P^4}\right), \end{aligned} \quad (68)$$

where k is the common amplitude of \mathbf{k}_1 and \mathbf{k}_2 and N_α is the first-order component of the gauge transformation, Eq. (61). Application of the chain rule to the contractions in (68) allows the Lyth–Rodríguez Taylor coefficients to be identified,

$$N_i \equiv \frac{\partial N}{\partial \phi_i} = N_\alpha \Gamma_{\alpha i}. \quad (69)$$

It follows that (68) is the standard result [5].

Three-point function.—Neglecting the initial three-point function $\mathcal{A}_{i'j'k'}$, the bispectrum can be computed by similar methods. There is an added complication from second-

order terms in the gauge transformation (59). Working

from (45) (or (53) and (55)) gives

$$\langle \zeta(\mathbf{k}_1)\zeta(\mathbf{k}_2)\zeta(\mathbf{k}_3) \rangle = (2\pi)^3 \delta(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) (N_\alpha \Gamma_{amn} + N_{\alpha\beta} \Gamma_{am} \Gamma_{\beta n}) N_i N_j \mathcal{S}_{mi} \mathcal{S}_{nk} \left(\frac{1}{k_1^3 k_2^3} + \frac{1}{k_1^3 k_3^3} + \frac{1}{k_2^3 k_3^3} \right) + \mathcal{O}\left(\frac{H^6}{M_p^6}\right). \quad (70)$$

We can make the identification

$$N_{ij} \equiv \frac{\partial^2 N}{\partial \phi_i \partial \phi_j} = N_\alpha \Gamma_{aij} + N_{\alpha\beta} \Gamma_{ai} \Gamma_{\beta j}, \quad (71)$$

where $N_{\alpha\beta}$ is the second-order term (66). The familiar approximation for the amplitude of the local mode, f_{NL} , follows immediately,

$$\begin{aligned} \frac{6}{5} f_{\text{NL}} &= \frac{N_{mn} N_j N_k \mathcal{S}_{mj} \mathcal{S}_{nk}}{(N_q N_r \mathcal{S}_{qr})^2} = \frac{N_\alpha \Gamma_{amn} N_j N_k \mathcal{S}_{mj} \mathcal{S}_{nk}}{(N_q N_r \mathcal{S}_{qr})^2} + \frac{N_{\alpha\beta} \Gamma_{am} \Gamma_{\beta n} N_j N_k \mathcal{S}_{mj} \mathcal{S}_{nk}}{(N_q N_r \mathcal{S}_{qr})^2} \\ &\equiv f_{\text{NL}}^\phi + f_{\text{NL}}^{\text{gauge}}. \end{aligned} \quad (72)$$

In the final step we have divided the contributions into an *intrinsic* term, f_{NL}^ϕ (which contains Γ_{amn}), and a *gauge contribution* $f_{\text{NL}}^{\text{gauge}}$ (which does not). The intrinsic term depends on the bispectrum of the fluctuations $\delta\phi_a$. Eq. (46) shows that it depends on $u_{\alpha\beta\gamma}$, and therefore has a memory of the *nonlinear* evolution of the connecting vectors along the trajectory. However, it has no dependence on the nonlinear part of the gauge transformation. *Vice versa*, the gauge term depends on the nonlinear part of the gauge transformation, and only on the *linear* evolution of the connecting vectors—that is, the Jacobi fields, in the guise of

the van Vleck matrix (16).

This separation was first made in Ref. [8], where it was shown that the gauge contribution dominated in a class of models known to generate large $|f_{\text{NL}}|$ [54]. We will sharpen this division slightly in Eqs. (84a)–(84b) below.

The outcome of this discussion is that f_{NL} could be computed efficiently by decomposing (72) into the component gauge transformations and Γ -symbols, which can be obtained using ordinary differential equations. An alternative approach is to work from the explicit formula (53), yielding

$$\langle \zeta(\mathbf{k}_1)\zeta(\mathbf{k}_2)\zeta(\mathbf{k}_3) \rangle = (2\pi)^3 \delta(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) (N_\alpha N_\beta N_\gamma \alpha_{\alpha|\beta\gamma} + N_{\alpha\beta} N_\gamma N_\delta \Sigma_{\alpha\gamma} \Sigma_{\beta\delta}) \left(\frac{1}{k_1^3 k_2^3} + \frac{1}{k_1^3 k_3^3} + \frac{1}{k_2^3 k_3^3} \right) + \mathcal{O}\left(\frac{H^6}{M_p^6}\right). \quad (73)$$

It then follows that

$$f_{\text{NL}}^\phi = \frac{5}{6} \frac{N_\alpha N_\beta N_\gamma \alpha_{\alpha|\beta\gamma}}{(N_\lambda N_\mu \Sigma_{\alpha\beta})^2} = \frac{5}{18} \frac{N_\alpha N_\beta N_\gamma \alpha_{\alpha\beta\gamma}}{(N_\lambda N_\mu \Sigma_{\alpha\beta})^2}. \quad (74)$$

In the final equality we have defined $\alpha_{\alpha\beta\gamma}$ by symmetrization,

$$\alpha_{\alpha\beta\gamma} \equiv \alpha_{\alpha|\beta\gamma} + \alpha_{\beta|\alpha\gamma} + \alpha_{\gamma|\alpha\beta}. \quad (75)$$

Note that this combination is not normalized to give weight unity. Eq. (54) shows that it obeys the transport equation (35) for the three-point function after dropping primes

on all indices. It was in this form that f_{NL} was quoted in Refs. [8], although the derivation was given in real space and is not the same as the one given here.

Gauge contribution.—There is some interest in isolating the gauge contribution to $|f_{\text{NL}}|$. As explained above, this is known to dominate in some models, including examples where large $|f_{\text{NL}}|$ is generated during a turn in field space [18, 54]. Comparison with (73) shows that it can be written

$$\frac{6}{5} f_{\text{NL}}^{\text{gauge}} = \frac{N_{\alpha\beta} N_\gamma N_\delta \Sigma_{\alpha\gamma} \Sigma_{\beta\delta}}{(N_\lambda N_\mu \Sigma_{\lambda\mu})^2}. \quad (76)$$

Combining (61) and (66) gives an explicit expression,

$$\frac{6}{5} f_{\text{NL}}^{\text{gauge}} = \frac{\eta}{2} - M_p v \left(\frac{\langle \sigma \delta \phi_a \rangle K_{\alpha\beta} \langle \delta \phi_\beta \sigma \rangle}{\langle \sigma \sigma \rangle^2} - 2 \frac{\langle \sigma \delta \phi_a \rangle \mathcal{D}_a \ln v}{\langle \sigma \sigma \rangle} \right), \quad (77)$$

where we have defined $\langle \sigma \sigma \rangle = \hat{n}_\alpha \hat{n}_\beta \Sigma_{\alpha\beta}$, and $\langle \sigma \delta \phi_a \rangle = \hat{n}_\beta \Sigma_{\alpha\beta}$. This expression is covariant under rotations of the

isocurvature plane. However, its form suggests a natural coordinate basis in which its content is more transparent. Since $K_{\alpha\beta}$ is symmetric, it can be diagonalized. Its eigenvectors form an orthonormal basis directed along the *principal curvature directions* of the fixed energy-density hypersurface in field space. We label these eigenvectors with an index \mathbf{m} and denote them $\zeta_{\alpha}^{\mathbf{m}}$, which can be considered as a vielbein. We can refer to the corresponding isocurvature directions as the *principal isocurvature modes*. The corresponding eigenvalues of the second fundamental form are the *principal curvatures* $k_{\mathbf{m}}$.

Next, we define a correlation coefficient $\rho_{\mathbf{m}}$ between the \mathbf{m}^{th} principal isocurvature mode and the adiabatic direction,

$$\zeta_{\alpha}^{\mathbf{m}} \langle \sigma \delta \phi_{\alpha} \rangle \equiv \langle \sigma \mathbf{m} \rangle = \rho_{\mathbf{m}} \langle \sigma \sigma \rangle^{1/2} \langle \mathbf{m} \mathbf{m} \rangle^{1/2}. \quad (78)$$

It is also useful to define analogues of the η -parameter for the isocurvature directions. It is a matter of convention how this is done. By analogy with our definition of η in the adiabatic direction we set

$$\eta_{\mathbf{m}} \equiv \frac{2M_{\text{P}}}{v} \zeta_{\alpha}^{\mathbf{m}} \mathcal{D}_{\alpha} \epsilon. \quad (79)$$

Unlike the adiabatic η -parameter, these isocurvature $\eta_{\mathbf{m}}$ -parameters need not be small even if slow-roll is an excellent approximation. In this basis we find

$$f_{\text{NL}}^{\text{gauge}} = \frac{\eta}{2} + \sum_{\mathbf{m}} \eta_{\mathbf{m}} \rho_{\mathbf{m}} \frac{\langle \mathbf{m} \mathbf{m} \rangle^{1/2}}{\langle \sigma \sigma \rangle^{1/2}} - M_{\text{P}} v \sum_{\mathbf{m}} k_{\mathbf{m}} \rho_{\mathbf{m}}^2 \frac{\langle \mathbf{m} \mathbf{m} \rangle}{\langle \sigma \sigma \rangle}. \quad (80)$$

As has been explained, these depend only on the Jacobi fields and geometrical quantities at the time of evaluation for f_{NL} . We have not displayed the \mathbf{k} -modes associated with these objects. Eq. (80) strictly applies for roughly comparable $|\mathbf{k}|$.

There are three contributions. First, there is the adiabatic η -parameter. As explained above this will almost always be negligible. Second, there is a weighted sum of $\eta_{\mathbf{m}}$ -parameters associated with the isocurvature directions. These may be individually large. Their contribution is suppressed by the correlation coefficient $\rho_{\mathbf{m}}$ between the adiabatic mode and fluctuations in the \mathbf{m}^{th} direction, and also by the “anisotropy factor” $(\langle \mathbf{m} \mathbf{m} \rangle / \langle \sigma \sigma \rangle)^{1/2}$ which measures their relative amplitude. Third, there is a weighted sum of the principal curvatures. These are weighted by the combination $\rho_{\mathbf{m}}^2 \langle \mathbf{m} \mathbf{m} \rangle / \langle \sigma \sigma \rangle$. Therefore, this term is typically dominant when the bundle has exaggerated extent in at least one isocurvature direction.

In a two-field model, Eq. (80) becomes especially simple. There is only one principal isocurvature mode, and it is orthogonal to the adiabatic direction. Also, the second fundamental form $K_{\alpha\beta}$ has a null eigenvector and therefore the principal curvature k is simply its trace. Comparison with (65) shows that

$$k = \text{tr} K_{\alpha\beta} = \frac{1}{M_{\text{P}} v} \left(\frac{M-1}{M} \theta - \hat{n}_{\alpha} \hat{n}_{\beta} \sigma_{\alpha\beta} \right). \quad (81)$$

As in §II, we have set M to be the dimension of field space.

Non-Gaussianity at the adiabatic limit.—There has been considerable interest in the fate of non-Gaussianity if an adiabatic limit is reached during slow-roll inflation. Meyers & Sivanandam [37] studied a class of models in which f_{NL} , g_{NL} and τ_{NL} decay to negligible values when all isocurvature modes decay, and argued that this behaviour is generic. However, explicit examples exist in which an observable value of f_{NL} persists even after all isocurvature modes are extinguished [18, 39, 55]. The separation of f_{NL} into intrinsic and gauge contributions allows us to shed further light on this issue.

At an adiabatic limit we expect $\langle \mathbf{m} \mathbf{m} \rangle \rightarrow 0$, and therefore Eq. (80) implies $f_{\text{NL}}^{\text{gauge}} \approx \eta/2$. The same conclusion can be obtained from (77) because any tensor projected onto the isocurvature plane (such as $K_{\alpha\beta}$ or \mathcal{D}_{α}) is orthogonal to $\Sigma_{\alpha\beta}$ in this limit. This is an advantage of the tensorial approach we have described, based on associating isocurvature modes with the tangent plane to surfaces of constant energy density in phase space.

One can also show that the intrinsic f_{NL} satisfies

$$f_{\text{NL}}^{\phi} = f_{\text{NL}}^{\phi, \text{AL}} + \frac{\eta_{\text{AL}}}{2} - \frac{\eta}{2}, \quad (82)$$

where ‘AL’ denotes evaluation just after the adiabatic limit is reached. In the language of §II E this may coincide with the onset of an inflow trajectory. We conclude that, at any subsequent time, f_{NL} has value

$$f_{\text{NL}} = f_{\text{NL}}^{\phi, \text{AL}} + \frac{\eta_{\text{AL}}}{2}, \quad (83)$$

which is constant as we expect. If the adiabatic limit is reached during slow-roll inflation, where η_{AL} must be small, this enables us to give a more precise formulation of Meyers & Sivanandam’s argument: if f_{NL} is large in the adiabatic limit, it must be because a large *intrinsic* three-point function is developed during the evolution. This is indeed the case in known examples where a large f_{NL} is reached in the “horizon-crossing approximation” [18, 39].

Eqs. (80), (82) and (83) also enable us to sharpen the division between “gauge” and “intrinsic” contributions. We define

$$f_{\text{NL}}^A = f_{\text{NL}}^{\phi} + \frac{\eta}{2} \quad (84a)$$

$$f_{\text{NL}}^B = f_{\text{NL}}^{\text{gauge}} - \frac{\eta}{2}. \quad (84b)$$

The advantage of this redefinition is that the *A*- and *B*-type contributions are constant at an adiabatic limit; indeed, f_{NL}^B is zero there because it captures only transient effects caused by the evolving isocurvature modes. However, when $|f_{\text{NL}}|$ is large the *A*- and *B*-type terms approximately correspond to the intrinsic and gauge f_{NL} .

This division is not unique, because a total derivative can always be added to the time integral in $\Gamma_{\alpha ij}$. However, the division in Eqs. (84a)–(84b) seems phenomenologically useful because all models (of which we are aware) which

generate large non-gaussianity do so in one of two ways: either f_{NL}^A becomes large at the adiabatic limit, or f_{NL}^B is large some time before the adiabatic limit is reached. As the following examples show, the underlying reason seems to be that the B -type term responds immediately to strong distortions of the shape of bundle, whereas the A -type term does not.

Example: Byrnes et al. model.—We illustrate Eqs. (77), (80) and (84a)–(84b) using examples drawn from the literature.

Consider the model $V = V_0 \phi^2 e^{-\lambda \chi^2}$ introduced by Byrnes et al. [54]. We follow their choices, setting $\lambda = 0.05 M_{\text{p}}^{-2}$ and fixing initial conditions $\phi = 16 M_{\text{p}}$ and $\chi = 0.001 M_{\text{p}}$. The first phase of evolution is descent from a ridge, during which a large spike in f_{NL} is generated by the gauge term. An interpretation of this contribution was given in Ref. [18].

In Fig. 6a we plot f_{NL} during the inflationary phase. For most of the evolution it is dominated by $f_{\text{NL}}^{\text{gauge}}$. In turn $f_{\text{NL}}^{\text{gauge}}$ is dominated by the extrinsic curvature term $K_{\alpha\beta}$. In Fig. 6b we plot the difference between the full f_{NL} and the $K_{\alpha\beta}$ -term, demonstrating explicitly that it is small.

In Figs. 6c–6f we plot the bundle parameters which determine the $K_{\alpha\beta}$ -term and the other contributions to $f_{\text{NL}}^{\text{gauge}}$. The correlation constant is initially zero but approaches -1 , making the curvature and isocurvature mode (anti-) correlated, as first discussed by Langlois [22]. The principal curvature k and isocurvature η -parameter exhibit only modest evolution over the entire range of e -folds. In comparison, the anisotropy factor $(\langle \mathbf{m}\mathbf{m} \rangle / \langle \sigma\sigma \rangle)^{1/2}$ grows dramatically. Its evolution is the dominant factor which determines the evolution of f_{NL} . A large f_{NL} arises because the ensemble of separate universes becomes highly anisotropic, with nearly twenty-five times as much power in the isocurvature direction as in the adiabatic direction. Evidently this must arise from a large contribution to the integrated shear in the propagator matrix.

Note that, although $f_{\text{NL}}^B \approx f_{\text{NL}}^{\text{gauge}}$ responds immediately to this strong anisotropy factor, there is no corresponding significant enhancement of the intrinsic three-point function.

In Fig. 6g and 6h we plot the bundle dilation, θ , and the focusing Θ . The dilation is always positive, so the bundle cross-section grows monotonically. Hence the total power in the isocurvature mode also grows monotonically. Evidently, the spike in f_{NL} is not due to the *total* isocurvature power, but to its relative growth compared with the adiabatic power. The large Θ implies that this model does not reach an adiabatic limit, and some other mechanism must be invoked to end inflation and determine the value of each observable. In Ref. [54] it was assumed that sudden destabilization of a waterfall field could play this role.

Example: axion quadratic model.—A similar phenomenon occurs in the axion–quadratic model discussed above. We plot the evolution of f_{NL} in Fig. 7a. It exhibits three distinct components. The first is a negative spike, generated by the axion rolling off its hilltop. The second is a smaller positive spike produced by the axion rolling into its min-

imum. These two spikes come from the gauge contribution to f_{NL} , as clearly shown in Fig. 7b. Fig. 7c shows that each spike is inherited from a spike in the anisotropy factor. This is consistent with the analysis of Elliston et al. [18], in which the spikes were interpreted as due to strong deformations in the shape of the bundle. In the present interpretation, the differing signs arise because the principal curvature changes sign in the intermediate evolution.

As for the Byrnes et al. model, the intrinsic term $f_{\text{NL}}^A \approx f_{\text{NL}}^\phi$ does not respond immediately to this strong anisotropy, growing only later on approach to the adiabatic limit. The anisotropy is due to a strong shearing effect arising near the turn from dominantly ϕ -evolution to dominantly χ -evolution. Near the deep negative spike in f_{NL} , there is an enhancement in the shear oriented parallel to the principal isocurvature mode. This enhances the fluctuations in the isocurvature direction.

The third feature is the flat plateau at late times, associated with the adiabatic limit. Fig. 7b shows that this comes from growth in the intrinsic term f_{NL}^ϕ ; see the discussion in Refs. [18, 39].

VI. SUMMARY

In this paper we have developed an analogy between inflationary perturbation theory and geometrical optics. Here we summarize the main steps in the discussion.

Background.—In inflationary perturbation theory, we are interested in following the statistical properties—as measured by the correlation functions—of an ensemble of spacetime regions. This ensemble can be constructed equally well within the separate universe picture or traditional cosmological perturbation theory.

The ensemble picks out a cloud of points in phase space. In the limit $k/aH \ll 0$, interactions between members of the ensemble are suppressed and each point moves along a phase space orbit of the unperturbed system. Therefore the ensemble traces out a narrowly-collimated “spray” or bundle of trajectories. Where slow-roll applies, the momenta are determined in terms of the fields and we can work in terms of a simplified flow on field space.

Optical quantities.—Geometrical properties of the bundle of trajectories can be used to describe its evolution and determine its statistical properties. The quantities of principal importance are obtained by decomposing the expansion tensor, yielding the dilation, shear and twist. These are well-known from the description of light rays in general relativity.

Jacobi fields and van Vleck matrix.—The dilation, shear and twist determine the evolution of *Jacobi fields*, which describe infinitesimal vectors connecting nearby trajectories. At any point in the flow, the van Vleck matrix aggregates the linearly independent Jacobi fields. The Jacobi fields themselves are measured from a fiducial trajectory, which can be thought of as the eikonal of geometrical optics. This

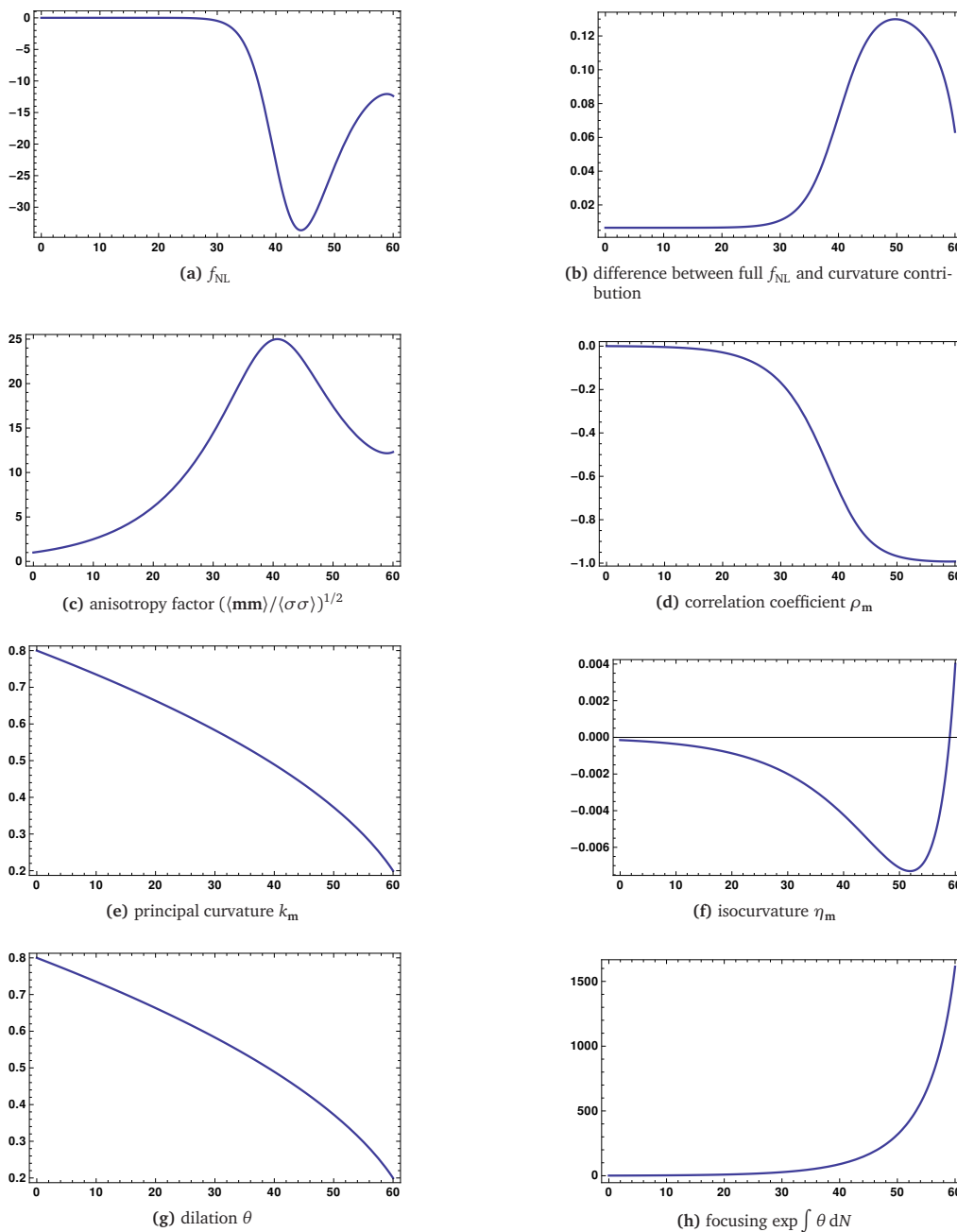


FIG. 6: Bundle parameters for the Byrnes et al. model $V = V_0 \phi^2 e^{-\lambda \chi^2}$. The initial conditions are $\phi = 16M_p$ and $\chi = 0.001M_p$, and $\lambda = 0.05M_p^{-2}$. All plots are against the e-folding number N , measured from horizon exit of the mode in question.

analogy is exact within the slow-roll approximation.

We have argued that different implementations of the separate universe assumption—such as the Lyth–Rodríguez Taylor expansion, or the transport equations of §III—can be thought of as different methods to compute the Jacobi fields, in the form of the van Vleck matrix (16). More generally, the same is true for all approaches to perturbation theory in the limit $k/aH \rightarrow 0$. The most familiar implementation of the separate universe assumption, the “ δN formalism” or Taylor expansion approach, follows from Ja-

cobi’s method of varying a solution with respect to its constants of integration. Conversely, the transport equations arise more naturally from Jacobi’s differential equation.

On approach to a caustic, some number of Jacobi fields decay. At an *adiabatic caustic*, defined in §IID, all but one of the Jacobi fields decay. The single remaining field represents fluctuations along the caustic. In inflation this mode is the adiabatic fluctuation. The other Jacobi fields represent isocurvature fluctuations between the spacetime regions which make up the ensemble. Therefore, focusing at

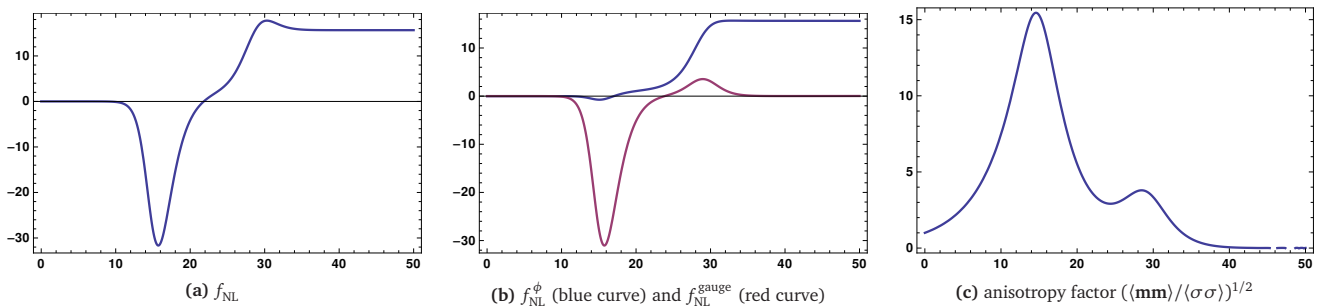


FIG. 7: Bundle parameters for the axion-quadratic model; for the potential and initial conditions, see Fig. 4. All plots are against the e-folding number N , measured from horizon exit of the mode in question.

an adiabatic caustic can be interpreted as decay of isocurvature modes, or approach to an adiabatic limit in the sense of Elliston et al. [18].

Transport equations.—The “ u -tensors” encode evolution of the connecting vector fields. We have argued that these tensors can be computed using either cosmological perturbation theory or the separate universe approximation. More generally, any formalism which can reproduce the \mathbf{k} -space deviation equation (31) will reproduce the correct correlation functions, because the u -tensors uniquely determine the transport equations. Therefore the u -tensors may be used as an objective way to compare competing formalisms.

The transport equations obtained in this way are generalizations of the transport equations previously introduced in Ref. [8]. Because they are expressed in terms of u -tensors, it follows that they can be integrated in terms of the Jacobi fields and their derivatives. Therefore the correlation functions can be expressed using the van Vleck matrix and its derivatives. (Technically it is the inverse of the van Vleck matrix which appears, in the form of the propagator matrix (15).)

In turn the van Vleck matrix can be expressed in terms of the integrated dilation, shear and twist. This makes it possible to diagnose regions where the flow may become adiabatic by tracking the behaviour of the focusing parameter Θ , defined in Eq. (19), and the behaviour of the shear and twist.

Working within the slow-roll approximation we have argued that $\Theta \gtrsim 1$ implies the presence of remaining isocurvature modes. To be compatible with experiment, these must almost certainly decay before the surface of last scattering. The consequent transfer of power into the adiabatic mode can change the value of ζ .

Flow equations.—The Jacobi fields yield a formal solution for each correlation function, analogous to the “line of sight” used to simplify integration of the Boltzmann equation in CMB codes.

This formal solution demonstrates explicitly that the transport equations reproduce the Taylor expansion algorithm of Lyth & Rodríguez. In doing so we also obtain explicit expressions for the Taylor coefficients Γ_{ai} and Γ_{aij} in

terms of integrals of the expansion tensor and its derivatives along the flow. Similar expressions had previously been obtained by Yokoyama et al. [7].

These explicit expressions can be manipulated to obtain a closed set of evolution equations for the Taylor coefficients. These are Eqs. (37a) and (50). Such equations are extremely helpful in practice, because it means the Taylor coefficients can be obtained without the challenging problem of extracting a variational derivative after numerical integration: without a sufficiently accurate integration algorithm, the small variation of interest can be swamped by numerical noise.

Transport of shape coefficients.—Even after obtaining the Taylor coefficients, it is necessary to extract coefficients for each type of momentum dependence (or “shape,” in inflationary terminology) which occurs in a correlation function. An alternative is to return to the full \mathbf{k} -space transport equations and derive evolution equations for these coefficients directly. The first nontrivial case is the three-point function, whose shape coefficients are determined by Eq. (54).

Gauge transformations.—Specializing to the slow-roll approximation, where the flow can be described in field space, ray-tracing techniques can be used to obtain the gauge transformation to ζ . In this way the gauge transformation is expressed using geometrical quantities in field space, rather than merely derivatives of the potential.

In models where a large f_{NL} is obtained from the gauge transformation, this gives a geometrical interpretation of its magnitude. The contributory factors are: (1) the η -parameters of the adiabatic and principal isocurvature modes; (2) the principal curvatures of uniform-density hypersurfaces in field space; (3) the correlation coefficient between the adiabatic fluctuations and the fluctuations in each principal isocurvature mode; and (4) an anisotropy factor which measures distortions in the cloud of field-space points representing the ensemble.

In two cases where a large, transient contribution to f_{NL} has been observed, we show this principally arises from a strong enhancement in the anisotropy factor.

Comparison with other geometrical formulations.—In common with all other approaches to the evolution of cor-

relation functions, the interpretation described in §§III–V is a reformulation of perturbation theory. All approaches carry the same physical content. Therefore, aside from practical considerations, the merit of each reformulation arises from the insight gained by emphasis on different structures.

The formulation we have given emphasizes the background phase space manifold, which encodes the structure of the theory in its geometry. This geometrical structure is mapped out by the behaviour of the trajectories flowing over it. Globally, this connection is made precise by the methods of Morse theory. Locally, it is encoded in the Jacobi fields whose role we have highlighted.

Attempts to reformulate perturbation theory in terms of geometrical objects have already attracted attention by various authors. Gordon et al. [56] and Nibbelink & van Tent [10] formulated perturbation theory for the two-point function in terms of the Frenet basis, which they called the “kinematical basis.” (See also Achúcarro et al. [13].) Peterson & Tegmark later extended this approach to the three-point function [12]. A Frenet basis can be defined for each trajectory, and the Frenet–Serret equation describes how this basis is transported along the trajectory. In Refs. [10, 12, 13, 56] these equations are used to describe transfer between the adiabatic and isocurvature modes.

In the Frenet formulation, the isocurvature modes are identified with the normal, binormal, \dots , vectors. In our formulation these modes arise from the eigenvectors of the extrinsic curvature, $K_{\alpha\beta}$, which we have described as the principal isocurvature modes. The tangent plane spanned by the Frenet normal, binormal, \dots , is the same as the subspace spanned by the eigenvalues of $K_{\alpha\beta}$, so the physical content of these formulations is the same. More generally, in our formulation the properties of the isocurvature modes are expressed using the familiar mathematical apparatus used to describe hypersurfaces—normal vectors, first and second fundamental forms, and so on.

In addition, we explicitly separate a “local” contribution to each ζ correlation function, arising from a gauge transformation and depending on the precise orientation of the Frenet basis, from the “integrated” contributions, obtained by solving the transport equations. Although it is clear that one can equally well express the integrated contributions in any suitable basis, it requires extra effort to rotate to the Frenet basis at each step in the integration. We feel it is preferable to express the evolution equations of perturbation theory in terms of the original basis on field space.

Future directions.—This formalism can be extended in several directions.

First, at some points in the discussion we specialized to the slow-roll approximation, to take advantage of certain simplifications—such as the twist-free and hypersurface-orthogonal character of flow. However, as we have presented it, the underlying formalism is independent of slow-roll. It can be used to evolve both field and momentum perturbations. This is desirable because future data from microwave background or galaxy surveys will be highly accurate, demanding commensurate accuracy in our theoret-

ical calculations.

Second, in this paper we have interpreted the decay of isocurvature modes, and approach to an “adiabatic” caustic. Our detailed discussion was restricted to field space. It should also be possible to study focusing and decay of isocurvature modes on the full phase space, providing a framework for the study of *kinetically* dominated scenarios, such as descent through the waterfall of hybrid inflation, where focusing may also occur.

Third, the existence of explicit expressions for the Taylor coefficients Γ_{ai} and Γ_{aij} may enable new analytic solutions to be found.

Finally, the entire formalism can be extended to higher n -point functions. The case of principal interest is the four-point function. In contrast to the three-point function, this requires two shape parameters which determine τ_{NL} and g_{NL} .

ACKNOWLEDGMENTS

DS was supported by the Science and Technology Facilities Council [grant numbers ST/F002858/1 and ST/I000976/1], and would like to thank the Harish-Chandra Research Institute, Allahabad, for their hospitality during the programme *Primordial Features and Non-Gaussianities* (December 2010) where an early version of this work was presented. JF was supported by the Science and Technology Facilities Council [grant number ST/1506029/1]. DJM was supported by the Science and Technology Facilities Council [grant number ST/H002855/1]. RHR was supported by Fundação para a Ciência e a Tecnologia through the grant SFRH/BD/35984/2007, and acknowledges a research studentship from the Cambridge Philosophical Society. We would like to thank Antony Lewis and Andrew Liddle for comments on a draft version of the manuscript.

Appendix A: Yokoyama et al. backwards formalism

In §IVB we presented integral formulae for the “ δN ” coefficients, Eqs. (49a)–(49b), and noted that essentially identical expressions had been presented by Yokoyama et al. [7]. (However, Yokoyama et al. obtained their results by very different means.) Since their work is closely related to our own in content and outlook, we take this opportunity to review and extend their results.

Their aim is to develop evolution or “transport” equations (in our terminology) for objects closely related to observation, such as the derivatives N_i —defined in section §V—and the f_{NL} parameter. They proceed as we do, first fixing a flat initial hypersurface. In our notation this is distinguished with lower case Roman indices. Unlike us, they also fix the final slice to be the precise time at which we wish to know the value of each observable quantity. In our notation this slice is labelled with Greek letters, and we

obtain the matrices Γ_{ai} and Γ_{aij} as a function of it. Observables can be obtained after evaluating these functions at the time of interest. Instead, Yokoyama et al. consider intermediate flat slices *between* the initial and final slices, and express their answers as a function of the intermediate time. As we now explain, observables are to be obtained by setting this intermediate slice equal to the *initial* hypersurface.

For convenience we extend our index notation, and label quantities evaluated on the intermediate slice with upper case Roman indices. Yokoyama et al. introduce the quantity

$$N_I = N_\alpha \Gamma_{\alpha I}, \quad (A1)$$

where N_α was defined in §V. N_I is the derivative of the number of e-folds between an intermediate flat hypersurface and the final uniform density hypersurface, with respect to the field values on the intermediate slicing. Yokoyama et al. introduce a further quantity $\Theta_I = \Gamma_{Ii} N_i$. Note that Θ_I is not to be confused with the focusing parameter Θ defined in the text, which is the exponential of the integrated dilation. N_I and Θ_I obey the autonomous transport equations

$$\frac{dN_I}{dN} = -u_{IJ} N_J, \quad (A2)$$

$$\frac{d\Theta_I}{dN} = u_{IJ} \Theta_J, \quad (A3)$$

Evaluating N_I at the final hypersurface gives N_α , which provides a boundary condition for the differential equation. One can then evolve *backwards* in time until we reach the initial slice. At this point N_I will equal N_i , which is the Taylor coefficient we set out to calculate. After this has been done, Θ_I can be evolved *forwards* from the initial hypersurface with boundary condition $\Theta_i = N_i$.

We describe this as the “backwards” formalism, to be contrasted with the “forwards” formalism we have described in the text.

The introduction of these quantities is ingenious. Employing Eq. (71) together with Eq. (49b) yields

$$N_{ij} = N_\alpha \Gamma_{\alpha i} \int_{N^*}^N \Gamma_{l\sigma}^{-1} u_{\sigma\beta\gamma} \Gamma_{\beta i} \Gamma_{\gamma j} dN' + N_{\alpha\beta} \Gamma_{\alpha i} \Gamma_{\beta j}, \quad (A4)$$

In turn this leads to

$$N_i N_{ij} N_j = \int_{N^*}^N N_I u_{IJK} \Theta_J \Theta_K dN' + \Theta_\alpha \Theta_\beta N_{\alpha\beta}. \quad (A5)$$

Therefore, f_{NL} can be evaluated with knowledge only of N_I , Θ_I and u_{IJK} .

In performing this calculation, Yokoyama et al. traded a three-index object (either Γ_{aij} or $\alpha_{\alpha|\beta\gamma}$, depending which formulation is in use) for two one-index objects, N_I and Θ_I . This involves fewer equations and therefore can be numerically advantageous.

Nevertheless, the backwards formalism has some disadvantages. First, because it computes only the Taylor coefficients, information about isocurvature modes is discarded. The evolution equations for $\Sigma_{\alpha\beta}$ and $\alpha_{\alpha|\beta\gamma}$, or Γ_{ai} and Γ_{aij} , allow the isocurvature modes to be retained.

Second, to obtain information about the time-evolution of any observable it is necessary to recalculate N_I and Θ_I with multiple *final* times. Although the method yields N_I , which is apparently related to the gauge transformation at an intermediate time, this is not quite correct. N_I is defined for a fixed *future* rather than past boundary condition, and therefore gives information about a range of scales at a fixed time of observation, rather than a fixed scale at a range of final times. The past-defined objects required for the latter are automatically provided by the forwards formalism, meaning that multiple integrations are not required.

If the time of observation is known then the backwards formalism gives an efficient means to treat multiple scales at once.

To extend the backwards formalism to the trispectrum, one needs to separate the observables τ_{NL} and g_{NL} . For this purpose N_{ij} and N_{ijk} themselves would be required. Therefore, given the potential utility of this method, we conclude by extending it to include a backwards evolution equation for N_{IJ} . As for the spectrum, this can be used to obtain information about f_{NL} and τ_{NL} over a range of scales at a fixed time of observation. It still requires the solution for only a two-index object. The transport equation for N_{IJ} can be shown to be

$$\frac{dN_{JK}}{dN} = -u_{IJK} N_I - u_{IJ} N_{IK} - u_{IK} N_{IJ}. \quad (A6)$$

This is to be solved backwards from the final hypersurface where N_{IJ} is equal to $N_{\alpha\beta}$.

-
- [1] E. Komatsu *et al.* (WMAP Collaboration), *Astrophys.J.Suppl.*, **192**, 18 (2011), arXiv:1001.4538 [astro-ph.CO].
 - [2] “The scientific programme of Planck,” (2006), arXiv:astro-ph/0604069 [astro-ph]; P. Ade *et al.* (Planck Collaboration), (2011), arXiv:1101.2022 [astro-ph.IM].
 - [3] V. Desjacques and U. Seljak, *Class.Quant.Grav.*, **27**, 124011

- (2010), arXiv:1003.5020 [astro-ph.CO].
- [4] A. A. Starobinsky, *JETP Lett.*, **42**, 152 (1985); D. H. Lyth, *Phys.Rev.*, **D31**, 1792 (1985); M. Sasaki and E. D. Stewart, *Prog.Theor.Phys.*, **95**, 71 (1996), arXiv:astro-ph/9507001; D. S. Salopek and J. R. Bond, *Phys. Rev.*, **D42**, 3936 (1990); M. Sasaki and T. Tanaka, *Prog.Theor.Phys.*, **99**, 763 (1998), arXiv:gr-qc/9801017 [gr-qc]; D. Wands, K. A. Malik, D. H.

- Lyth, and A. R. Liddle, Phys.Rev., **D62**, 043527 (2000), arXiv:astro-ph/0003278 [astro-ph].
- [5] D. H. Lyth and Y. Rodríguez, Phys.Rev.Lett., **95**, 121302 (2005), arXiv:astro-ph/0504045.
- [6] G. I. Rigopoulos and E. P. S. Shellard, JCAP, **0510**, 006 (2005), arXiv:astro-ph/0405185; G. I. Rigopoulos, E. P. S. Shellard, and B. J. W. van Tent, Phys.Rev., **D73**, 083521 (2006), arXiv:astro-ph/0504508.
- [7] S. Yokoyama, T. Suyama, and T. Tanaka, JCAP, **0707**, 013 (2007), arXiv:0705.3178 [astro-ph]; Phys.Rev., **D77**, 083511 (2008), arXiv:0711.2920 [astro-ph]; JCAP, **0902**, 012 (2009), arXiv:0810.3053 [astro-ph].
- [8] D. J. Mulryne, D. Seery, and D. Wesley, JCAP, **1001**, 024 (2010), arXiv:0909.2256 [astro-ph.CO]; **1104**, 030 (2011), arXiv:1008.3159 [astro-ph.CO].
- [9] L. Amendola, C. Gordon, D. Wands, and M. Sasaki, Phys.Rev.Lett., **88**, 211302 (2002), arXiv:astro-ph/0107089 [astro-ph].
- [10] S. Groot Nibbelink and B. van Tent, Class.Quant.Grav., **19**, 613 (2002), arXiv:hep-ph/0107272 [hep-ph].
- [11] Z. Lalak, D. Langlois, S. Pokorski, and K. Turzyński, JCAP, **0707**, 014 (2007), arXiv:0704.0212 [hep-th].
- [12] C. M. Peterson and M. Tegmark, Phys.Rev., **D83**, 023522 (2011), arXiv:1005.4056 [astro-ph.CO]; **D84**, 023520 (2011), arXiv:1011.6675 [astro-ph.CO]; (2011), arXiv:1111.0927 [astro-ph.CO].
- [13] A. Achúcarro, J.-O. Gong, S. Hardeman, G. A. Palma, and S. P. Patil, JCAP, **1101**, 030 (2011), arXiv:1010.3693 [hep-ph].
- [14] A. Avgoustidis, S. Cremonini, A.-C. Davis, R. H. Ribeiro, K. Turzyński, et al., JCAP, **1202**, 038 (2012), arXiv:1110.4081 [astro-ph.CO].
- [15] J.-L. Lehnens and S. Renaux-Petel, Phys.Rev., **D80**, 063503 (2009), arXiv:0906.0530 [hep-th].
- [16] C. Ringeval, Lect.Notes Phys., **738**, 243 (2008), arXiv:astro-ph/0703486 [astro-ph]; J. Martin and C. Ringeval, JCAP, **0608**, 009 (2006), arXiv:astro-ph/0605367 [astro-ph].
- [17] I. Huston and K. A. Malik, JCAP, **0909**, 019 (2009), arXiv:0907.2917 [astro-ph.CO]; **1110**, 029 (2011), arXiv:1103.0912 [astro-ph.CO].
- [18] J. Elliston, D. J. Mulryne, D. Seery, and R. Tavakol, JCAP, **1111**, 005 (2011), arXiv:1106.2153 [astro-ph.CO].
- [19] L. Boubekeur and D. Lyth, Phys.Rev., **D73**, 021301 (2006), arXiv:astro-ph/0504046 [astro-ph]; D. H. Lyth, JCAP, **0606**, 015 (2006), arXiv:astro-ph/0602285 [astro-ph]; D. Seery, Class.Quant.Grav., **27**, 124005 (2010), arXiv:1005.1649 [astro-ph.CO].
- [20] D. Polarski and A. A. Starobinsky, Phys.Rev., **D50**, 6123 (1994), arXiv:astro-ph/9404061 [astro-ph].
- [21] J. García-Bellido and D. Wands, Phys.Rev., **D53**, 5437 (1996), arXiv:astro-ph/9511029 [astro-ph]; D. Wands and J. García Bellido, Helv.Phys.Acta, **69**, 211 (1996), arXiv:astro-ph/9608042 [astro-ph].
- [22] D. Langlois, Phys.Rev., **D59**, 123512 (1999), arXiv:astro-ph/9906080 [astro-ph].
- [23] S. Hawking and G. Ellis, *The Large scale structure of space-time* (Cambridge University Press, 1973).
- [24] F. de Felice and C. Clarke, *Relativity on curved manifolds* (Cambridge University Press, 1990).
- [25] R. H. Ribeiro and D. Seery, JCAP, **1110**, 027 (2011), arXiv:1108.3839 [astro-ph.CO]; T. Battefeld and J. Grieb, **1112**, 003 (2011), arXiv:1110.1369 [astro-ph.CO].
- [26] D. Seery and J. E. Lidsey, JCAP, **0509**, 011 (2005), arXiv:astro-ph/0506056 [astro-ph].
- [27] R. Sachs, Proc.Roy.Soc.Lond., **A264**, 309 (1961); R. Penrose, in *Perspectives in Geometry and Relativity*, edited by B. Hoffman (Indiana University Press, Bloomington, Indiana, 1966) p. 259.
- [28] H. Fried, *Green's functions and ordered exponentials* (Cambridge University Press, 2002) ISBN 9780521443906.
- [29] G. Rigopoulos, E. Shellard, and B. van Tent, Phys.Rev., **D73**, 083522 (2006), arXiv:astro-ph/0506704 [astro-ph]; **D76**, 083512 (2007), arXiv:astro-ph/0511041 [astro-ph].
- [30] F. Vernizzi and D. Wands, JCAP, **0605**, 019 (2006), arXiv:astro-ph/0603799 [astro-ph].
- [31] T. Battefeld and R. Easther, JCAP, **0703**, 020 (2007), arXiv:astro-ph/0610296 [astro-ph]; D. Seery and J. E. Lidsey, **0701**, 008 (2007), arXiv:astro-ph/0611034 [astro-ph].
- [32] M. Visser, Phys.Rev., **D47**, 2395 (1993), arXiv:hep-th/9303020 [hep-th].
- [33] C. DeWitt-Morette, Annals Phys., **97**, 367 (1976); C. DeWitt-Morette and T. Zhang, Phys.Rev., **D28**, 2503 (1983); C. DeWitt-Morette, T. Zhang, and B. Nelson, **D28**, 2526 (1983).
- [34] D. H. Lyth, K. A. Malik, and M. Sasaki, JCAP, **0505**, 004 (2005), arXiv:astro-ph/0411220 [astro-ph].
- [35] G. Rigopoulos and E. Shellard, Phys.Rev., **D68**, 123518 (2003), arXiv:astro-ph/0306620 [astro-ph].
- [36] S. Weinberg, Phys.Rev., **D78**, 123521 (2008), arXiv:0808.2909 [hep-th]; **D79**, 043504 (2009), arXiv:0810.2831 [hep-ph].
- [37] J. Meyers and N. Sivanandam, Phys.Rev., **D83**, 103517 (2011), arXiv:1011.4934 [astro-ph.CO]; **D84**, 063522 (2011), arXiv:1104.5238 [astro-ph.CO].
- [38] S. Dimopoulos, S. Kachru, J. McGreevy, and J. G. Wacker, JCAP, **0808**, 003 (2008), arXiv:hep-th/0507205 [hep-th].
- [39] S. A. Kim and A. R. Liddle, Phys.Rev., **D74**, 023513 (2006), arXiv:astro-ph/0605604 [astro-ph]; S. A. Kim, A. R. Liddle, and D. Seery, Phys.Rev.Lett., **105**, 181302 (2010), arXiv:1005.4410 [astro-ph.CO]; Phys.Rev., **D85**, 023532 (2012), arXiv:1108.2944 [astro-ph.CO].
- [40] J. L. Cardy, *Scaling and renormalization in statistical physics* (Cambridge University Press, 1996).
- [41] J. Silk and M. S. Turner, Phys.Rev., **D35**, 419 (1987).
- [42] L. Alabidi and D. H. Lyth, JCAP, **0605**, 016 (2006), arXiv:astro-ph/0510441 [astro-ph].
- [43] E. Calzetta and B.-L. Hu, Phys.Rev., **D55**, 3536 (1997), arXiv:hep-th/9603164 [hep-th].
- [44] E. Calzetta and B. Hu, Phys.Rev., **D61**, 025012 (2000), arXiv:hep-ph/9903291 [hep-ph].
- [45] P. R. Jarnhus and M. S. Sloth, JCAP, **0802**, 013 (2008), arXiv:0709.2708 [hep-th].
- [46] C. W. Gardiner, *Handbook of stochastic methods: for physics, chemistry and the natural sciences*, Springer Series in Synergetics, 13 (Springer, 2002) ISBN 3540616349.
- [47] J. Smidt, A. Amblard, C. T. Byrnes, A. Cooray, A. Heavens, et al., Phys.Rev., **D81**, 123007 (2010), arXiv:1004.1409 [astro-ph.CO]; J. Fergusson, D. Regan, and E. Shellard, (2010), arXiv:1012.6039 [astro-ph.CO].
- [48] D. H. Lyth and D. Seery, Phys.Lett., **B662**, 309 (2008), arXiv:astro-ph/0607647 [astro-ph].
- [49] D. Seery, JCAP, **0711**, 025 (2007), arXiv:0707.3377 [astro-ph]; **0802**, 006 (2008), arXiv:0707.3378 [astro-ph].
- [50] C. Lam, J.Math.Phys., **39**, 5543 (1998), arXiv:hep-th/9804181 [hep-th].
- [51] M. Dias and D. Seery, Phys.Rev., **D85**, 043519 (2012),

- arXiv:1111.6544 [astro-ph.CO].
- [52] D. H. Lyth and I. Zaballa, JCAP **0510**, 005 (2005), arXiv:astro-ph/0507608 [astro-ph].
 - [53] K. A. Malik and D. Wands, Phys.Rept., **475**, 1 (2009), arXiv:0809.4944 [astro-ph].
 - [54] C. T. Byrnes, K.-Y. Choi, and L. M. Hall, JCAP **0810**, 008 (2008), arXiv:0807.1101 [astro-ph]; **0902**, 017 (2009), arXiv:0812.0807 [astro-ph].
 - [55] D. Mulryne, S. Orani, and A. Rajantie, Phys.Rev., **D84**, 123527 (2011), arXiv:1107.4739 [hep-th].
 - [56] C. Gordon, D. Wands, B. A. Bassett, and R. Maartens, Phys.Rev., **D63**, 023506 (2001), arXiv:astro-ph/0009131 [astro-ph].